ROBUST ESTIMATION OF SYSTOLIC TIME INTERVALS USING BALLISTOCARDIOGRAM AND SEISMOCARDIOGRAM SIGNALS

A Dissertation
Presented to
The Academic Faculty

by

Hazar Ashouri

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in
Electrical and Computer Engineering

Georgia Institute of Technology
May 2018

ROBUST ESTIMATION OF SYSTOLIC TIME INTERVALS USING BALLISTOCARDIOGRAM AND SEISMOCARDIOGRAM SIGNALS

Approved by:

Dr. Omer T. Inan, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Amit Shah
Rollins School of Public Health
Emory University

Dr. Pamela T. Bhatti
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Mozziyar Etemadi
School of Biomedical Engineering
Northwestern University

Dr. David V. Anderson
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: Dec 08, 2017

To my parents, Maan Ashouri and Hala Haidar,

and in loving memory of my friend, Leya Hojeij

ACKNOWLEDGEMENTS

Many people were essential to the completion of this thesis either on a professional or a personal level, or both. I would like to dedicate this section to acknowledging the contributions of these people.

I would like to start by thanking my advisor, Professor Omer Inan. I got introduced to Prof. Inan almost a year after starting my graduate studies in Georgia Tech. Within a month, it was clear to me that I would not find a better advisor. Little did I know back then that this would be one of the best decisions I have ever made. Honestly, I do not think I would have been able to continue with a Ph.D. had I not been working with him. He was always there to provide guidance and support. I remember going into many of our meetings thinking that I have hit a dead-end, but I have always left those meetings with a new prospective and excitement to try new ideas. He was also very supportive of my professional plans and even helped me accomplish them more than once. I am very grateful for him and will never forget the role he played in my life.

I would also like to thank Dr. Pamela Bhatti, Dr. David Anderson, Dr. Mozziyar Etemadi and Dr. Amit Shah for taking the time to be in my dissertation committee. They provided me with very insightful feedback regarding my research and the completion of my Ph.D.

I was fortunate to work alongside smart and friendly labmates. I would like to thank all my collegues in Inan Research Lab for their help when I needed it and their constuctive feedback on my proposal and dissertation: Dr. Sinan Hersek, Dr. Abdul Qadir Javaid, Dr. Hakan Toreyin, Dr. Ozan Bicen, Dr. Maziyar Baran Pouyan, Oludotun Ode, Andrew

Carek, Caitlin Teague, Nick Bolus, Samer Mabrouk, Mobashir Shandhi, Nil Gurel, Hyeon Ki Jeong, Daniel Whittingslow, Beren Semiz, Venu Ganti, Heywon Jung, and Lara Orlandic. Specifically, I would like to thank Dr. Abdul Qadir Javaid with whom I collaborated heavily in the early stages of my research and Dr. Sinan Hersek who helped me greatly throughout my research and mostly in its later stages. I would also like to thank Andrew Carek who made data collection fun during our collaboration, Nick Bolus who took the time to teach me how to use the 3D printer, and Lara Orlandic who helped me find and collect data from human subjects on my first study. Finally, I would like to thank Dr. Hakan Toreyin and Oludotun Ode who, in addition to Dr. Javaid, were the first people I work with in the lab and published with them the first conference paper related to my research.

I came to Atlanta having one acquaintance and no friends. Today, I have a group of friends that I consider family. They turned Atlanta into home to me. My friends, Leya Hojeij, Norah Essali, Sinan Hersek, Nour Deen Alnoury, Amr Shehadeh, Ayda Sawaf , Sara Al Moukadem and Leen Rmeih, I could not have done it without you. Those people were my support system and each one of them changed me into a better person.

Leya is the purest soul I met in my life. The unconditional love she had for every one of us cannot be matched. She always went the extra mile to make people around her happy even at her own expense. Leya passed away almost two years ago, but I know that she is watching over me and I want her to know that she was, and will always be, part of me.

Norah is simply my soul mate. She is the sister that I never had. She knows exactly what to say in any situation and has a big heart that never judges but always supports. I

cannot recall a happy moment for me in Atlanta that Norah was not part of or a sad time that Norah did not help me get through. Even after she moved to Augusta , to pursue her residency in psychiatry, our friendship never changed and it never will.

There are certain people that change the course of your life, Sinan was one of those people for me. If it was not for Sinan, I would not have met my advisor, prof. Omer Inan, and perhaps, would not have been able to go through the Ph.D. He is my lab mate, my best friend, my lunch buddy and my neighbor. I will always be grateful to him.

On the long tiring work days, the phone call from Nour at the end of the day was always something to look forward to. It was always helpful for me to bounce ideas with him and hear his point of view, which was different than mine most of the times. I value his opinion greatly and he definitely influenced many of my decisions and actions. Our friendship has and will continue to withstand any challenge and I am lucky to have that.

Amr was the only person that I knew when I came to Atlanta as we both went to the same highschool in Damascus, Syria. If it was not for him, I might have not met the other amazing people mentioned here. He quickly went from being just an acquaintance to becoming one of the closest people to me. Whenever I am down and in need for a good laugh, I know that he is the one I should go to.

We usually look up to people older than us, but I look up to Ayda even though she is two years younger than me. I met Ayda in Georgia Tech and we became close friends in a short period of time. She graduated and moved back to Dubai two years ago, yet our friendship only got closer despite the distance. I can say, without hesitation, that Ayda,

through her advice and support, is the person that helped me the most in achieving my professional life goals post PhD. For that, I will always be indebted to her.

It was unfortunate that I did not get to meet Sara during my undergraduate studies, even though we both went to the same school in Beirut, Lebanon. However, my luck turned when I got a second chance at that by meeting her here in Atlanta, only to wish I had known her from before. Her uplifting attitude, warm smile, and encouraging words were always a source of comfort whenever the PhD life seemed to get daunting.

I cannot recall a time that I needed Leen's help and she was not there for me. She is someone I can always count on. Leen taught me to be kind to people yet fearless in going after what I want. Knowing that she is in Atlanta, long term, gives me a sense of comfort and relief, that no matter what, I will always have family in here.

Lastly, but most importantly, I would like to express my deepest gratitude for my family; my dad, Maan Ashouri, my mom, Hala Haidar, and my brother, Fahmi Ashouri. Fahmi has and will always be the closest person to my heart. I am proud of him and I look forward to all the amazing things he will accomplish. My parents always supported me at every turn, from backing me up with my decision to do my undergraduate studies outside Syria, at the American University of Beirut, to encouraging me to come to the States to pursue my Ph.D., even though it meant being away from them, again. If it was not for their unconditional love, trust, sacrifice, and support, I would not even dream of being where I am today. Nothing I ever say or do can do them justice. I feel extremely grateful and so fortunate to have such incredible parents to look up to. You two are my role models and I wish I can, one day, be half the person either of you are.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

LIST OF TABLES xi

LIST OF FIGURES xii

SUMMARY xvi

CHAPTER 1. Introduction 1

1.1 Motivation 1

1.2 Major Contributions of this Work 8

1.3 Thesis Organization 10

CHAPTER 2. Unobtrusive Estimation of Cardiac Contractility and Stroke Volume Changes Using Ballistocardiography measurements on a high bandwidth force plate 12

2.1 Introduction 12

2.2 Methods 12

2.2.1 Protocol 12

2.2.2 Hardware Design 13

2.2.3 Data Processing 14

2.2.4 Feature Extraction 15

2.2.5 Correlation at Rest 16

2.2.6 Correlation during Exercise Recovery 17

2.2.7 Estimating Relative Changes in Stroke Volume 17

2.2.8 Signal-to-Noise Ratio Calculations 18

2.3 Results and Discussions 19

2.3.1 RI and PEP Correlation for Scale and Force Plate BCG during Rest 19

2.3.2 RI and PEP Correlation for Head-to-Foot Force Plate BCG during Recovery 20

2.3.3 Stroke Volume Estimation from Head-to-Foot Force Plate BCG during Recovery 21

2.3.4 Signal-to-Noise Ratio Comparison for Scale BCG and Force Plate Head-to-Foot BCG 24

2.4 Conclusions 27

CHAPTER 3. Classification of Wearable Seismocardiogram Sensor Positioning for Robust Estimation of the Pre-Ejection Period in Unsupervised Settings 29

3.1 Introduction 29
3.2 Methods 30
3.2.1 Protocol 31
3.2.2 Hardware Design 31
3.2.3 Data Processing 32
3.2.4 Feature Extraction 32
3.2.5 Correlation During Rest and Exercise Recovery 34
3.2.6 Classification of Different Positions of Accelerometers 35
3.3 Results and Discussions 39
3.3.1 PEP and R-AO Correlation Results for SCG from Different Sensor Positions 39
3.3.2 Classification Results 41
3.4 Conclusions 44

CHAPTER 4. Universal Pre-Ejection Period Estimation using Seismocardiography: Quantifying the Effects of Sensor Placement and Regression Algorithms 46
4.1 Introduction 46
4.2 Methods 47
4.2.1 Protocol 47
4.2.2 Hardware and Data processing 48
4.2.3 Feature Extraction 50
4.2.4 Regression Model 51
4.2.5 SCG Sensor Location Comparison 54
4.2.6 Rigid vs. Flexible Interfacing Material Between the SCG Sensor and Sternum 55
4.2.7 Feature Importance Evaluation 55
4.2.8 Comparing Different Regression Techniques 56
4.2.9 Evaluating the Effect of XGBoost Hyperparameters 58
4.2.10 Statistical Analysis 59
4.3 Results 60
4.3.1 Comparison of Different Sensor Locations 60
4.3.2 Comparison of Different Sensor Interfacing Material 62
4.3.3 Comparison of Different Regressors 64
4.3.4 Effect of XGBoost Hyperparameters 65

4.4 Discussion **66**

4.5 Conclusions and Future Work **69**

CHAPTER 5. Conclusion and future direction **71**

5.1 Conclusion **71**

5.2 Future Directions **72**

REFERENCES **74**

LIST OF TABLES

Table 1	Comparison of several studies on non-invasive cardiac output measurement for exercise and exercise recovery (healthy subjects).	23
Table 2	Per subject errors in ΔSV estimation.	24
Table 3	Per subject SNR calculations for scale and force plate BCG.	24
Table 4	Best features in terms of information gain	36
Table 5	The average correlation coefficient (r) obtained from linear, quadratic, and cubic regression for the five different positions across all subjects.	40
Table 6	Average absolute PEP estimation errors (ms) for each position	41
Table 7	RMSE (ms) for PEP estimates from SCG signals measured from sensors placed in different locations (Str, PMI, LC, RC)	62

LIST OF FIGURES

Figure 1	(a) A healthy person's heart. (b) The heart of an HF patient. It can be observed that the heart makes up for the its inability to pump enough blood by enlarging, which causes the body to retain fluids. Adapted from [3].	1
Figure 2	Changes in heart rate, right ventricular systolic pressure and pulmonary artery diastolic pressure before and after hospitalization. Adapted from [14].	3
Figure 3	The cardiac cycle and systolic time intervals.	4
Figure 4	(a) Echo measuring system. (b) The placement of the eight electrodes needed to measure the ICG signals.	6
Figure 5	(a) Adapted from [44]. A bed equipped with BCG measuring sensors. (b) A weighing scale with BCG measuring circuit. (c) A force plate used for measuring BCG. (d) Adapted from [45]. A tri-axial SCG measurement system. (e) An ultra-low noise tri-axial instrumentation accelerometer for SCG measurement. (f) Adapted from [46]. A wearable patch that measures ECG and SCG signals.	7
Figure 6	Sensing ecosystem for heart failure monitoring	8
Figure 7	A block diagram of the experimental setup.	14
Figure 8	A 5-seconds time trace showing ECG, ICG, head-to-foot force plate BCG.	15
Figure 9	Ensemble averaged traces of ECG, ICG, scale BCG, and head-to-foot force plate BCG with the characteristic points and features.	16
Figure 10	(a) Linear regression fit for both scale and force plate head-to-foot BCG RI-interval vs PEP among all subjects. (b) Bland Altman plot for scale and force plate linear prediction models of PEP. The blue line is the 95% confidence range of the PEP estimations from the scale BCG RI interval while the red line is the 95% confidence range of the PEP estimations from the force plate BCG RI interval.	20
Figure 11	(a) Linear regression fit for head-to-foot force plate BCG vs PEP during recovery among all 17 subjects.	21
Figure 12	Estimated stroke volume percent changes from head-to-foot force plate BCG compared to calculated stroke volume percent changes from reference ICG.	23

Figure 13 (a) a subject standing on the force plate based system for BCG signal measurement with feet d cm apart. (b) Average SNR across subjects for different stance widths. The lowest SNR was found with the feet placed together (0 cm), and the highest SNR was for a fairly wide stance (22.9 cm). 27

Figure 14 a) Adapted from [58]. An illustration of the experimental setup. (b) A 450 ms portion of one ensemble averaged beat of an SCG signal collected from the accelerometer placed on the mid-sternal position vs SCG signals collected from the other accelerometers at rest. It can be observed that the waveforms are very different in their shapes and result in different estimations of the aortic valve opening (AO) and aortic valve closure (AC) points. 32

Figure 15 (a) The variation in AO and peak-to-peak amplitude of the DV SCG signal collected from mid-sternum as the physiological state of the subject changes between rest and exercise recovery. It can be noted that right after exercise the AO time decreases while the peak-to-peak amplitude increases only to stabilize back to baseline value as time passes. The five rectangles represent the AO and peak-to-peak values of the five waveforms shown in part (d) for the different physiological states. (b) A 450 ms portion of an ensemble average of DV SCG beats collected from the sternum during rest and different states during exercise recovery. We can observe that the waveforms look similar for a time shift and an increase in amplitude after exercise. 33

Figure 16 An illustration of some of the top features in the time and frequency domain. 37

Figure 17 The training and testing sets for training and testing the classifier on one of the subjects. The shaded areas are features that are excluded from the training set and included in the testing sets when trying to detect the sensor position for the subject. 38

Figure 18 (a) The best feature in terms of info gain (Peak-to-Peak in the 1st 250 ms) plotted against the 3rd feature (RMS) for above sternum SCG compared to mid-sternum SCG in one of the test sets (b) The 2nd feature (Standard deviation) plotted against the 4th feature (1st peak of the PSD) for above sternum SCG compared to mid-sternum SCG in one of the test sets. It is clear in both (a) and (b) that those features result in two distinct clusters for these two positions which indicates that the position of the sensor (accelerometer) can be accurately detected. 43

Figure 19 ICG and dorsoventral SCG ensemble averaged traces (n = 5 heartbeats) obtained with the sensor on the sternum for three 47

different subjects. The ICG B-points and SCG AO-points are marked with red circles, and there is a substantial time difference between the two corresponding points for the three subjects: in two cases, the ICG B-point occurs first, and in the third case the SCG AO-point occurs first.

Figure 20 taneously. (b) Five beat ensemble averaged traces of ECG, ICG and mid-sternum dorsoventral SCG heartbeats. The ECG R-peak is used as a reference point for beat segmentation, the B-point of the ICG is used to detect aortic valve opening and the R-B interval is used as the ground truth PEP. Peak timing locations and width are extracted from the SCG signal as shown. (c) After extracting the features from the head-to-foot and dorsoventral axes of the SCG signals from all locations, a regression model is used to obtain PEP estimates from the features obtained from a single location, multiple combination of locations, one axis, and both axes. RMSE between the ground truth PEP and every estimate is calculated and the optimal location / combination of location and axes is determined. 49

Figure 21 (a) Different methods of interfacing the ADXL354 accelerometer with the sternum. (b) SCG signals obtained from each of the different interfacing materials. 56

Figure 22 (a). RMSE from PEP estimated from features obtained from accelerometers placed on the sternum (Str), point of maximal impulse (PMI), below the left clavicle (LC), and below the right clavicle (RC) for both the dorsoventral axis (z-axis) and head-to-foot and dorsoventral axes combined (z+x axes). (b) RMSE from PEP estimated from features obtained from the best performing combination of accelerometer locations. It can be observed that adding more sensors does not substantially reduce the error obtained using one sensor below the left or right clavicle. 61

Figure 23 Ranking of best 15 features obtained from the combination of sensors and axis that rendered the lowest RMSE (Str+LC axis z). 62

Figure 24 RMSE from PEP estimated from accelerometers placed on the sternum with different interfacing techniques: directly on the sternum (Str), in the middle of a silicone rubber sheet placed along on the sternum (fstr) and two accelerometers coupled by a rigid plastic mold and placed on the upper sternum (US) and lower sternum (LS) 63

Figure 25 Comparing RMSE for PEP estimates obtained using ensemble regression models vs. linear regression models on features obtained 65

from SCG signals that performed best with XGBoost (i.e., LC+sternum z-axis).

Figure 26 (a) RMSE for PEP estimates obtained using XGBoost on features obtained from LC+sternum z-axis while varying the learning rate parameter (b) RMSE for PEP estimates obtained using XGBoost on features obtained from LC+sternum z-axis while varying the column sample parameter. 65

Figure 27 Spearman correlation between the top 3 features from signals from every location and the I-wave, J-peak, and K-point of scale BCG. There was almost no correlation between sternum and PMI features with BCG features while good correlation was obtained between RC and LC features with BCG features and specifically the I-wave and J-peak. (highlighted with the red boxes). 67

SUMMARY

Heart failure (HF) is a cardiovascular disease characterized by a weakened myocardium, where the heart is unable to supply sufficient blood to the tissues and organs. It affects nearly 6 million Americans with annual health care costs in the United States being almost $31 billion, half of which are due to hospitalizations. Therefore, there has been great interest in developing novel methods for monitoring patients with HF at home to potentially reduce hospitalizations. This monitoring can be achieved by sensing the mechanical aspects of the cardiovascular health, specifically the estimation of systolic time intervals, such as the pre-ejection period (PEP), outside of clinical settings. PEP is a surrogate measure of myocardial contractility; an increased PEP signifies decreased contractility and hence, a weakened myocardium.

Unfortunately, existing technologies for non-invasive PEP measurements are expensive, obtrusive, and usually require a trained medical professional. Recent research has shown that ballistocardiogram signals (BCG), which measure the reactionary forces of the body in response to cardiac ejection of blood in the vasculature, and seismocardiogram signals (SCG), which measure local chest vibrations associated with the heart and blood movement, can be used to estimate systolic time intervals, including PEP. BCG and SCG can be measured using unobtrusive and inexpensive equipment, such as weighing scales with added circuitry and low noise accelerometers. However, there are gaps in the research that must be bridged before BCG and SCG can be reliably used in unsupervised settings to monitor cardiovascular health.

This work focuses on bridging some of these existing gaps. New hardware to measure BCG was explored by using a high bandwidth force plate and further signal processing techniques which allowed for significant improvement in absolute measurements of PEP, and measurements of changes in stroke volume, over current state-of-the-art technology. Additionally, since SCG signals measure local vibrations, the relationship between sensor placement and the morphology of the signals was investigated. This was done by designing a robust algorithm that distinguishes between changes in morphology resulting from placement and changes resulting from physical activities, and consequently, detects misplacement of the SCG sensors allowing for robust PEP monitoring in unsupervised settings. Moreover, different placements and interfaces of SCG sensors, on the upper body, were explored to identify the ideal position/ combination of positions. This showed, for the first time, that better PEP estimates can be obtained by placements different than what is currently used in research. Finally, a universal ensemble regression model, that uses multiple features to estimate PEP from SCG signals, is presented in this work. This algorithm bypasses the lack of a well-defined standard to detect the aortic valve (AO) opening from SCG, resulting from the signals morphology being affected by age, sex and heart condition.

The algorithms and methods presented in this work, pave the way towards enabling accurate and robust monitoring of the mechanical health of the cardiovascular system. Using these unobtrusive, inexpensive systems for HF patients would increase their quality of life by reducing rehospitalization, and therefore reducing the associated healthcare costs.

CHAPTER 1. INTRODUCTION

1.1 Motivation

Cardiovascular diseases (CVD) are the leading cause of death in the United States, according to the American Heart Association. A particularly prevalent CVD is Heart Failure (HF), which is caused by weakened myocardial contractility which leads to the inability of the heart to supply the cells with the amount of blood needed for normal function [1]. This results in shortness of breath, fatigue and swelling in the ankles, feet, legs, abdomen, and veins in the neck. Those symptoms result from the fluid build-up in the body due to the enlargement of the heart chambers as the ventricular muscles become weaker [2]. Figure 1(a) shows a healthy heart while Figure 1(b) shows the heart of an HF patient.

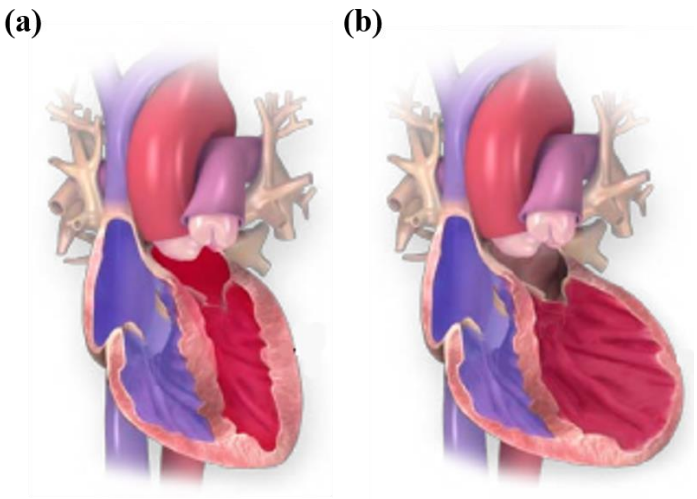

Figure 1. (a) A healthy person's heart. (b) The heart of an HF patient. It can be observed that the heart makes up for the its inability to pump enough blood by enlarging, which causes the body to retain fluids. Adapted from [3].

Today, 5.7 million Americans are living with HF and 1 in 9 deaths results directly from it [4, 5]. It costs the nation a total of $30.7 billion in medical costs. Projections show

that the number of Americans diagnosed with HF will rise to 8 million by 2030, a 46% increase from today's estimates [6]. Moreover, it is estimated that the medical cost of HF will increase by 127% to a total of $69.7 billion [6]. Another staggering statistic shows that approximately 27% of patients with HF are readmitted to the hospital within 30 days of their discharge with this number increasing to more than 50% readmission rate within 6 months [7, 8]. Intuitively, reducing rehospitalization rates would provide HF patients with a better quality of life and would reduce the medical cost associated with the disease.

There has been a significant increase in the efforts aiming to reduce hospital readmission rates and make care proactive through continuous home monitoring of HF patients [9]. It was thought that monitoring the weight of HF patients can help in detecting worsening conditions and adjusting the medication accordingly since decompensation leads to fluid build-up in the body which would lead to an increase in weight [10]. However, a study based on a large population of HF patients showed that telemonitoring patients based on their weight fails to provide additional benefit over usual care and does not lead to a decreased readmission rate [11]. This can be attributed to the fact that a patient's weight can increase due to higher caloric intake rather than from gaining fluid weight, which can lead to false positives. Similarly, the patient may be gaining fluid weight but at the same time losing body fat or skeletal muscles which would prevent the increase of the overall body weight [12, 13].

Later research showed that most HF readmissions are a result of an increase of the ventricular filling pressure which can be observed up to 6 days prior to symptoms onset as shown in Figure 2 [14]. However, studies have shown that the only efficient way in monitoring that in a home setting is through the use of implantable hemodynamic

monitoring (IHM) devices [15]. These devices have a very high cost with CardioMEMS IHM costing around $25,000.

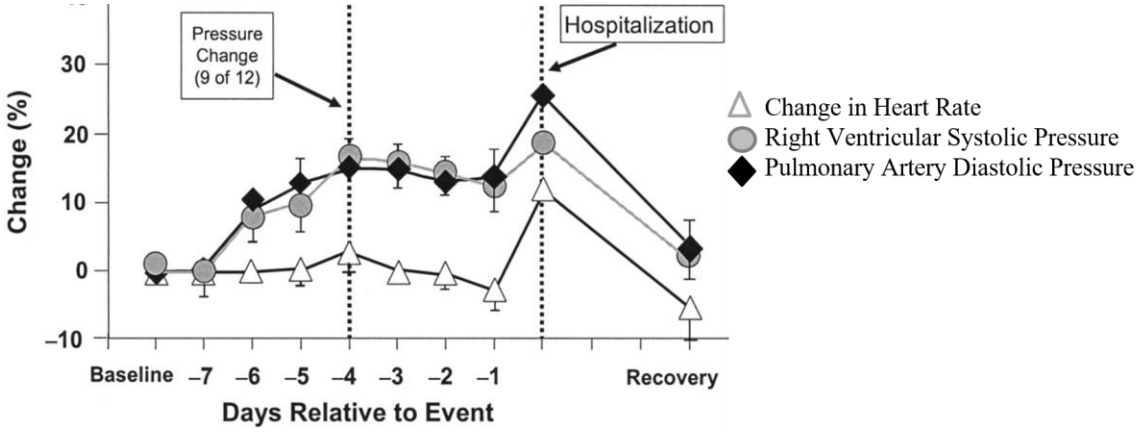

Figure 2. Changes in heart rate, right ventricular systolic pressure and pulmonary artery diastolic pressure before and after hospitalization. Adapted from [14].

However, there are physiological features other than ventricular filling pressure that can potentially allow the prediction of exacerbations in HF patients; those are the systolic time intervals. Each cardiac cycle is divided into filling and ejection periods, also known as diastole and systole. The time spent by the heart in each of these periods is an important marker of the mechanical function of the heart and vasculature, and thus provides insight into the health of the cardiovascular system. Therefore, accurate measurement of the timings of these periods, and in particular systole, is important in monitoring patients with cardiovascular diseases, including HF patients. Systole is divided into two time intervals known as the systolic time intervals (STI): pre-ejection period (PEP), which is the period of isovolumetric contraction, and left ventricular ejection time (LVET), which is the period of systolic ejection [16]. The cardiac cycle and the systolic time intervals are shown in Figure 3.

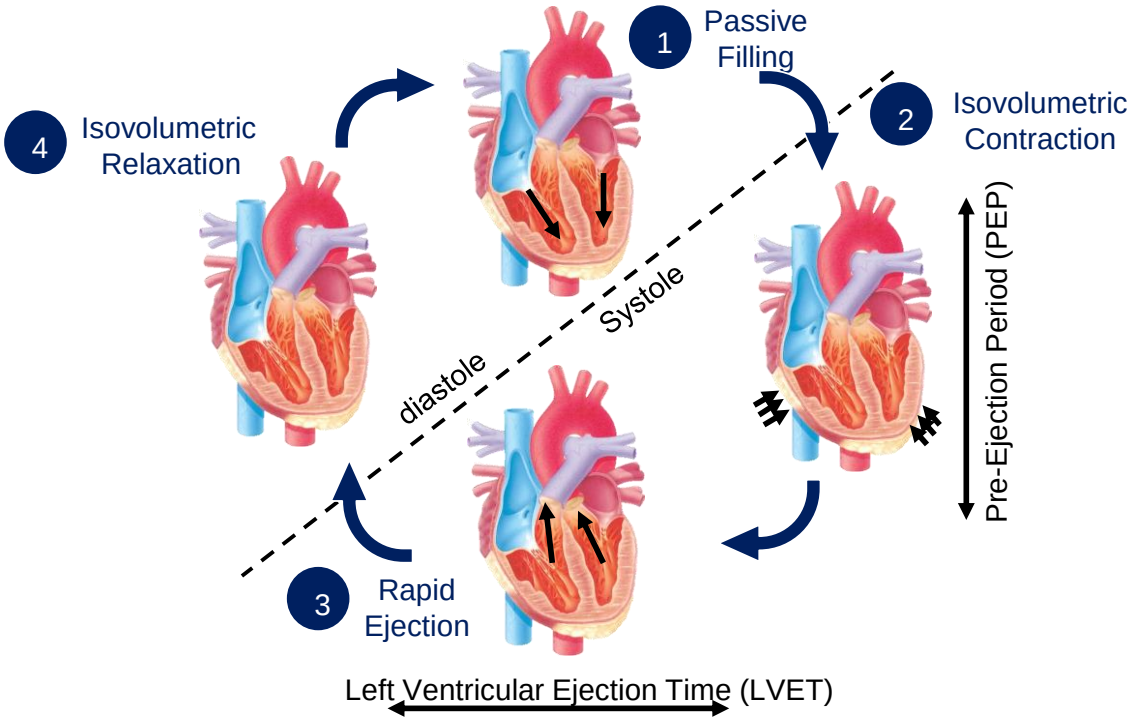

Figure 3. The cardiac cycle and systolic time intervals. After [17].

PEP is the time elapsed between the electrical depolarization of the ventricular muscle and the ensuing opening of the aortic valve [18], and is a surrogate measure of myocardial contractility [19]. Specifically, increased PEP signifies decreased contractility, and vice versa, since a weakened myocardium will have lowered maximal rate of left ventricular pressure increase (dP/dt_{max}) during isovolumetric systole. Therefore, an increased PEP can potentially be a sign of decompensation in HF patients. Hence, to realize out-of-clinic monitoring of HF patients and transition from reactive to proactive care, it is important to accurately measure STIs non-invasively, unobtrusively, and without the need for a health care professional.

Current non-invasive wearable technologies for cardiovascular monitoring are limited to cardiac electrophysiology measurements through electrocardiogram (ECG) [20, 21] and peripheral blood volume pulse through photoplethysmogram (PPG) [22, 23].

Although these technologies are noninvasive (wristband, smartwatches, adhesive patches, etc..[24]) and inexpensive, they can only provide information related to heart rate and rhythm but no information related to the mechanical aspects of the cardiovascular function and STIs such as PEP [25].

Existing technologies for noninvasive PEP measurement are obtrusive, and impractical for large scale use in HF monitoring. The reference standard for noninvasive PEP measurement is echocardiography, which uses ultrasound. An ultrasound transducer is placed over the chest and the sound waves bounce off the heart and echo back to the transducer. The waves are then changed into pictures viewed on a video monitor. The apparatus used for that is large and expensive and has be operated by trained medical staff [26, 27]. Impedance cardiography (ICG) was validated against echo as a reliable way of measuring PEP [28], however, it is also difficult to use in out-of-clinic settings. ICG is obtrusive to measure as it requires a trained medical professional to apply eight electrodes to the subject and operate the measuring device [29, 30]. Moreover, although some studies have shown that serial ICG measurements can predict the risk of acute decompensation [31], the requirement of frequent office visits and sensitivity to electrode positioning make this technology inapplicable on a larger scale [32]. Recently, researchers have developed a wearable system to measure ICG through textile integration [26], however, a more convenient way that would allow users to wear their standard clothing and still be able to obtain continuous PEP measurements is still needed. Conventional Echo and ICG measuring methods are shown in Figure 4.

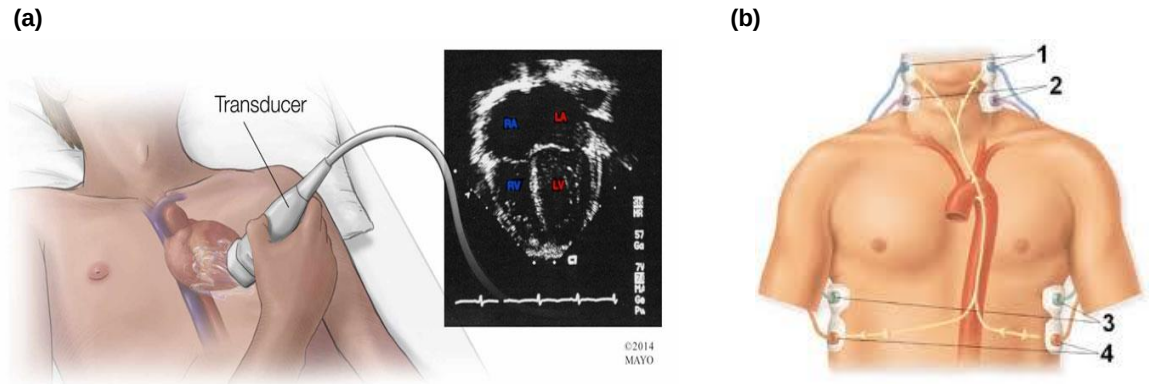

Figure 4. (a) Echo measuring system. (b) The placement of the eight electrodes needed to measure the ICG signals.

Recent studies have shown that the ballistocardiogram (BCG) signal, a measurement of the recoil forces of the body in response to the ejection of blood from the heart and movement of the blood through the vasculature [33, 34], and seismocardiogram (SCG) signals, representing the local chest vibrations associated with heart and blood movement, can be used to estimate systolic time intervals including PEP [33, 35].

BCG was first discovered by J. W. Gordon in 1877 when he observed that the needle of a scale, on which a person was standing, was synchronously vibrating with each heartbeat. He attributed that movement to the upward recoil force of the body as the blood is ejected downward into the aorta which he compared to the recoil force resulting from the propelling of a bullet from a gun [36]. However, it was not until 1939 that this signal was measured by Starr using a table-like instrument and he suggested that it can be used to estimate changes in cardiac output [34]. Despite that, BCG research did not gain momentum until the 21st century as developing small and low-cost sensors to measure it was not feasible before then. BCG can be measured using instrumented beds [37] and

chairs [38], weighing scales [39] and force plates [40]. Figure 5 (a), (b), and (c) shows some of the BCG acquisition hardware.

Likewise, and in contrast to the ICG, SCG measurements do not require any medical professional to administer the test, and in terms of hardware can be rather compact and unobtrusive as it is usually measured using a high resolution miniature accelerometer [41-43]. Acceleration can be measured in the 3-dimensions (head-to-foot, dorsoventral, and lateral), but most studies focus on the dorsoventral component. Since SCG measure local vibrations, the signal is impacted by the exact location of the sensor on the chest. Figure 5 (d), (e), and (f) shows some of SCG acquisition hardware.

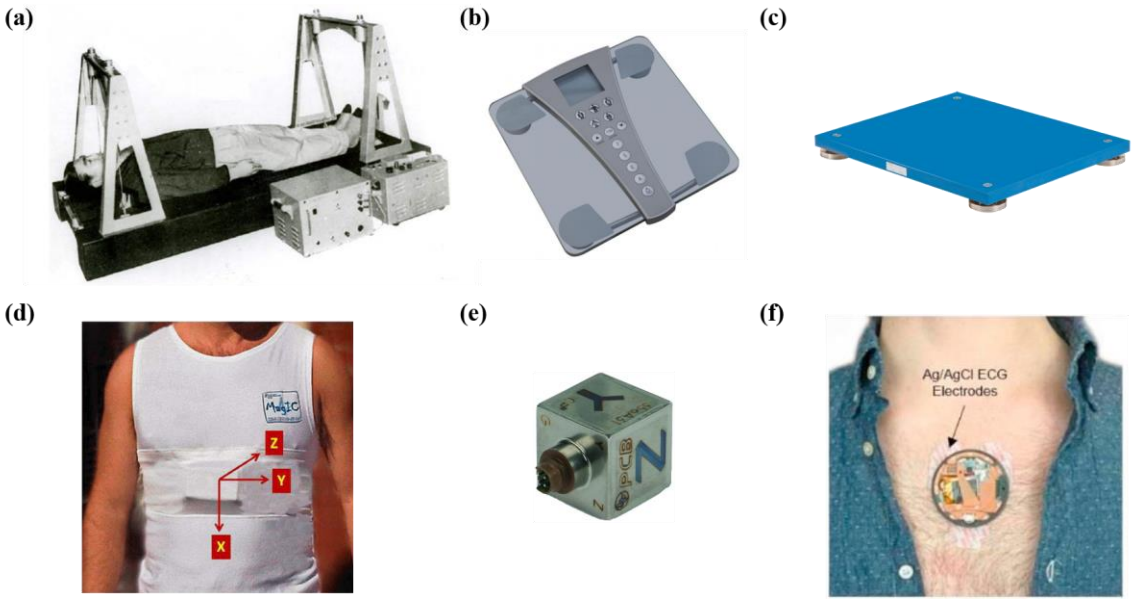

Figure 5. (a) Adapted from [44]. A bed equipped with BCG measuring sensors. (b) A weighing scale with BCG measuring circuit. (c) A force plate used for measuring BCG. (d) Adapted from [45]. A tri-axial SCG measurement system. (e) An ultra-low noise tri-axial instrumentation accelerometer for SCG measurement. (f) Adapted from [46]. A wearable patch that measures ECG and SCG signals.

The aim of this research is to pave the way for continuous out of clinic monitoring of patients with HF using systolic time intervals estimations obtained from SCG and BCG.

The end goal would be for the patient to take an SCG wearable device, or a BCG scale home, use the device on a daily basis while measurements are recorded and pushed into a cloud. STI estimation algorithms are then run on the data and the outputs are accessed by physicians, who will notify the patient if an adjustment of treatment is needed, allowing for titration of care. Figure 6 shows the vision for a sensing ecosystem for heart failure monitoring.

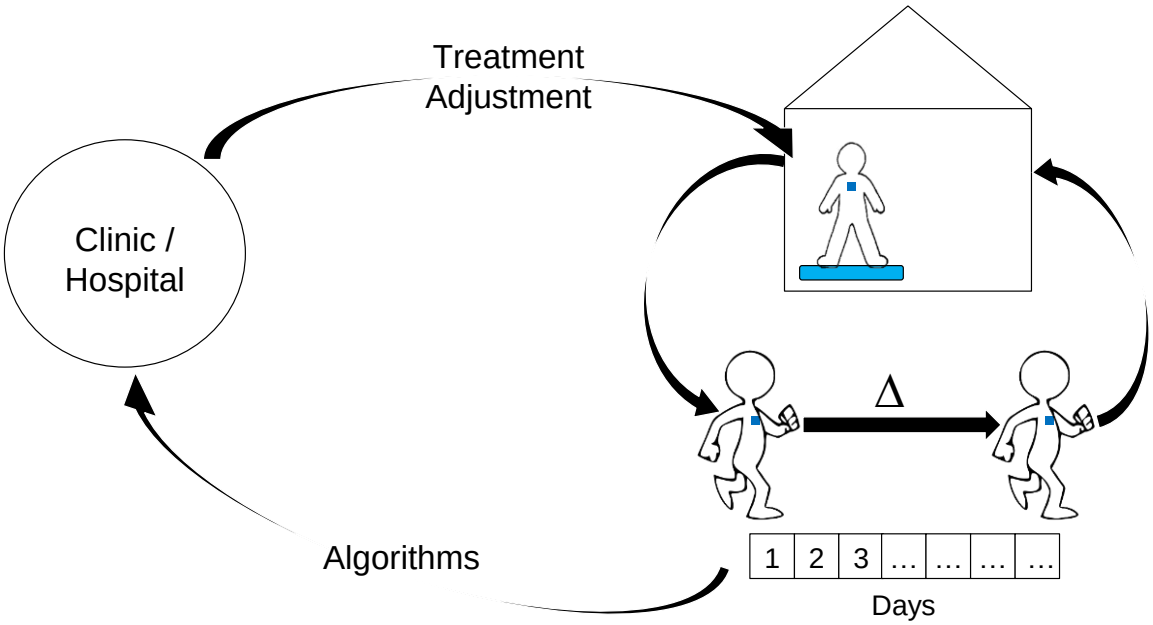

Figure 6. Sensing ecosystem for heart failure monitoring

1.2 Major Contributions of this Work

There has been a considerable amount of work done on the use of BCG and SCG for measuring mechanical parameters of the cardiovascular function. However, there are gaps in the research that must be bridged before BCG and SCG signals can be reliably used in unsupervised settings to monitor the cardiovascular health.

To measure BCG signals, a weighing scale with a strain gauge bridge and analog amplifier circuit is used. Weighing scales are usually designed to measure DC body weights, and using them to measure BCG signals, which are AC changes, could impose limitations. Additionally, the uneven group delay in the passband can potentially distort the BCG waveform and reduce the accuracy of STI measurements.

SCG signals result from local chest vibrations, therefore, the placement of the sensor affects the shape of the measured signals. This change in morphology, if not detected, affects the accuracy of the STI measurements. Hence, it is important to study SCG sensor placement, whether that is to identify the ideal positioning, or to control for misplacements. Additionally, the morphology of the SCG signal and its fiduciary points changes significantly with gender, age, and cardiac conditions, making it difficult to develop a standard method to detect these fiduciary points needed for PEP estimation, limiting SCG to subject-specific calibration and monitoring.

With these limitations in mind, the main contributions of this work are given below:

1. Demonstrated significant improvement in STI estimation accuracy compared to state-of-the-art BCG instrumentation using a high bandwidth force plate, showing that it is possible to use BCG signals to measure absolute values of PEP rather than only relative changes.
2. Demonstrated for the first time that placement based changes in SCG morphology can be differentiated from physiological changes and devised an algorithm to automatically detect sensor misplacement in unsupervised settings, improving PEP measurement accuracy by more than 10 ms.

3. Identified the optimal placement and interfacing method for SCG sensors on the upper body. This resulted in showing, for the first time, that improved PEP estimates (~30% improvement) can be obtained by placements different than what is currently used in research.
4. Devised a universal ensemble regression model, that uses multiple features extracted from the SCG signals, to estimate PEP. This model bypasses the lack of a well-defined standard to detect the aortic valve (AO) opening from SCG, and is robust to slight changes in placement since it is global, rather than subject-specific.

1.3 Thesis Organization

The rest of this thesis is organized as follows: Chapter 2 focuses on BCG signals as the modality for non-invasive monitoring of the mechanical health of the cardiovascular system while both chapter 3 and chapter 4 focus on SCG wearable sensors for enabling continuous monitoring in the context of normal stressors encountered by the person during their everyday activities.

Chapter 2 investigates acquiring BCG signals with a high bandwidth force plate compared to the modified weighing scale in terms of STI estimation accuracy. The increased bandwidth is then leveraged to estimate per-subject stroke volume changes. In Chapter 3, we explore the effect of sensor placement on the shape of SCG signals, prove that inaccurate placement can lead to inaccurate estimates of STIs, and present an algorithm that detects sensor misplacement in an unsupervised setting. Further placements are examined in chapter 5 and sensor fusion techniques, at the feature level, are used to identify

the ideal combination of sensor locations. Additionally, a non-linear universal model, to estimate PEP from multiple features in the SCG signal, is developed using ensemble regression techniques. Finally, Chapter 5 concludes the work in this thesis and provides direction for future research.

CHAPTER 2. UNOBTRUSIVE ESTIMATION OF CARDIAC CONTRACTILITY AND STROKE VOLUME CHANGES USING BALLISTOCARDIOGRAPHY MEASUREMENTS ON A HIGH BANDWIDTH FORCE PLATE

2.1 Introduction

The modified weighing scale used to measure BCG in [47] can be modeled as a second-order mechanical system. The spring constant and associated mechanical bandwidth for a range of bodyweights for typical adults were characterized (>15 Hz bandwidth for bodyweights of 150 kg or less) [47]. This mechanical bandwidth, while sufficient for not attenuating any of the frequency components of interest for BCG recordings, may potentially distort the BCG waveform and reduce the accuracy of timing interval measurements due to uneven group delay in the pass-band.

Therefore, we compare the modified weighing scale to a high bandwidth, reference standard force plate to evaluate the accuracy in PEP estimation. We also estimate per subject ΔSV (i.e., relative changes in stroke volume) using the force plate BCG and verify the accuracy against simultaneously measured reference standard ICG measurements.

2.2 Methods

2.2.1 Protocol

The study was conducted under a protocol reviewed and approved by the Georgia Institute of Technology Institutional Review Board (IRB). All subjects provided written consent before experimentation. Data were collected from 17 healthy subjects (Gender: 10 males, 7 females; Age: 23.6 ± 4.5 years, Height: 172.8± 9.9 cm, Weight: 70.7±11.3 kg).

Each subject was asked to stand still in an upright position for 60 seconds on each of the force plate and scale for baseline measurements. After that, each subject performed a stepping exercise for 60 seconds after which they were asked to stand on the force plate as still as possible for 5 minutes to monitor their full recovery. Half of the subjects stood on the scale first during baseline measurement while the other half stood on the force plate first to account for any bias due to the order in which the subject stood on the force plate versus the scale. In addition to the BCG, we measured the ECG and ICG. The ECG was used as timing reference to ensemble average the BCG and ICG signals. The B-point of the ICG was used as a reference method for detection of the opening of the aortic valve [16, 48]; it was also used for reference ΔSV calculations.

2.2.2 Hardware Design

Two different instruments were used to measure the BCG: a modified electronic weighing scale and a high bandwidth multi-component force plate. The modified weighing scale (BC534, Tanita, Tokyo, Japan) was developed in previous work using an analog amplifier and a strain gauge bridge [47]. The multi-component force plate (Type 9260AA6, Kistler Instrument Corp, NY, USA) has a bandwidth in excess of 200 Hz and sufficient resolution to enable BCG measurements in all three axes. The force plate has four, three-component force sensors. Each of the four sensors has three pairs of quartz plates, one sensitive to pressure in the z-direction (head-to-foot) and two to shear in the x and y-directions (dorsoventral and lateral respectively). Out of the 12 output signals, two of the shear forces that have the same line of action can be paralleled, so the outputs of the force plate are eight signals instead of 12. Since we only process head-to-foot BCG in this work, we accessed each of the four signals in the z-direction separately, passed each one of them

as an input to an amplifier and filter circuit, and then the outputs of these four components were added using an adder circuit. The summed, amplified and filtered head-to-foot BCG signal was outputted into the data acquisition system (MP150, BIOPAC System Inc., Goleta, CA). The ECG and ICG signals were measured using BN-RSPEC and BN-NICO wireless modules respectively (BIOPAC Systems, Inc., Goleta, CA) then transmitted to the MP150. All signals were sampled at 2 kHz. Figure 7 shows a block diagram of the experimental setup and Figure 8 shows a time trace of ECG, ICG, and head-to foot BCG.

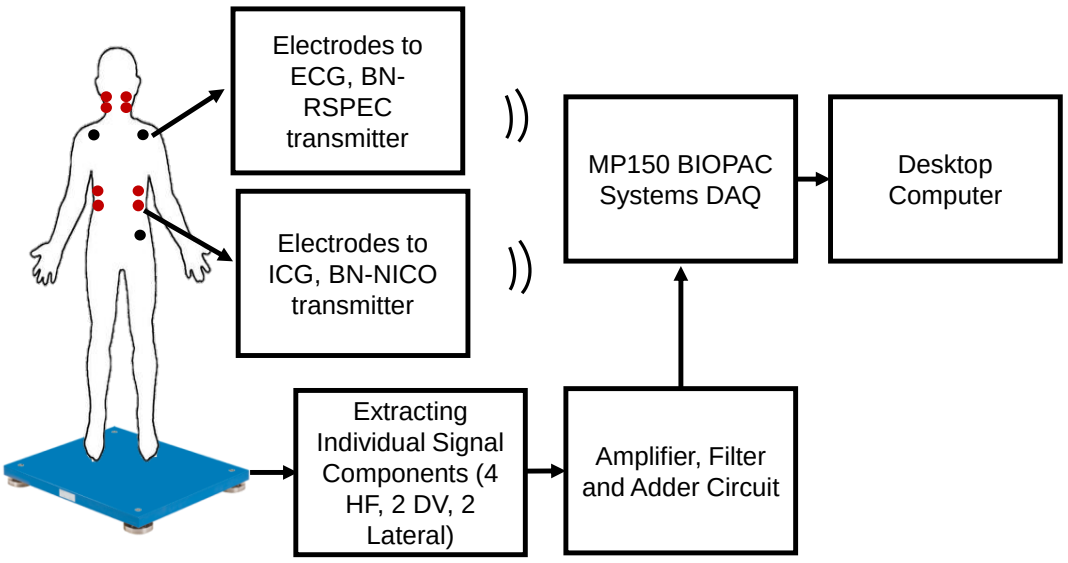

Figure 7. A block diagram of the experimental setup.

2.2.3 Data Processing

The BCG, ICG, and ECG signals were filtered with finite impulse response (FIR) Kaiser window band-pass filters with cut-off frequencies as follows: 0.8-35 for ICG and ECG, 0.5-20 for head-to-foot BCG. The ECG R-peaks were detected using an automated peak detection algorithm and verified manually. This was done by computing a constant threshold of 50% of the maximum amplitude of the absolute value of the band-pass filtered

ECG signal for each subject; local maxima, greater in amplitude than this threshold, were then located automatically and annotated as R-waves. The minimum R-R interval was calculated for every subject and the detected R-peaks were used as a fiduciary to segment ICG and BCG signals into individual heart beats, each with a window length equal to the minimum R-R interval. These extracted heartbeats were then averaged to obtain ensemble averaged traces with reduced noise.

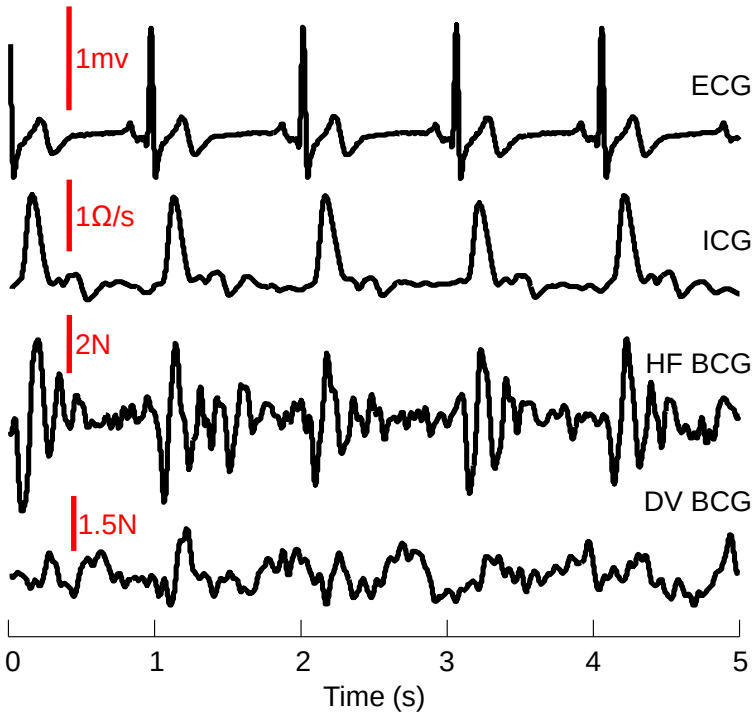

Figure 8. A 5-seconds time trace showing ECG, ICG, head-to-foot force plate BCG.

2.2.4 Feature Extraction

The I-wave in the weighing scale and force plate BCG signals was obtained by detecting the minima before the global maxima in the first 200 ms portion of the corresponding BCG ensemble average. The J-peak and its amplitude (from the reference threshold, i.e. 0) were detected as the global maxima of the force plate and scale BCG ensemble averages. The B-point of the ICG was detected as the global maxima of the first

derivative of the ICG [30]. The X-point was obtained by detecting the minima after the global maxima of the ICG. The amplitude of the maxima of the ICG was also detected. Figure 9 shows ensemble averages of ECG, ICG, scale BCG and head-to-foot force plate BCG signals with the extracted features.

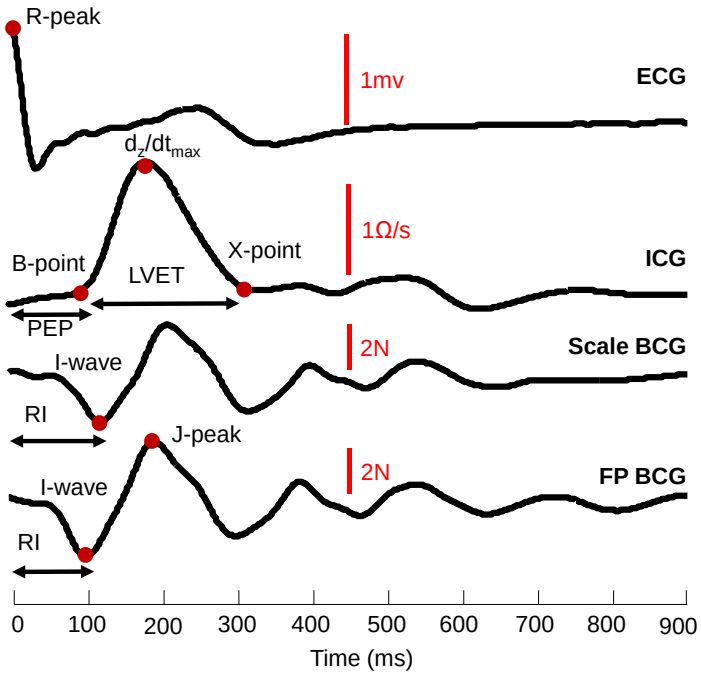

Figure 9. Ensemble averaged traces of ECG, ICG, scale BCG, and head-to-foot force plate BCG with the characteristic points and features.

2.2.5 Correlation at Rest

The 60 seconds baseline recording of scale BCG, force plate head-to-foot BCG, and ICG were divided into 12 5-second windows and an ensemble average was obtained for each of the 12 windows. For each of those ensemble averages, the RI duration was calculated (the timing at which the I-wave occurs since the ensemble averaging is done with respect to the ECG R-peak) and PEP was calculated (the timing at which the B-point – opening of the aortic valve – occurs). Although PEP is defined as the time elapsed between the Q-point in the ECG and the B-point in the ICG, detecting the Q-point in the

ECG is not always as robust as detecting the R-point. Hence, in our analysis we used the R-peak in ECG as a reference and estimated PEP as the RB interval since we are most interested in relative changes in PEP rather than absolute measures, and the QR interval is typically consistent from beat-to-beat A correlation analysis was then performed between RI-intervals in both scale and force plate and the corresponding PEP values (from ICG) among all 17 subjects.

2.2.6 Correlation during Exercise Recovery

The 5 minutes after exercise recovery recording of head-to-foot force plate BCG and ICG were divided into 15 20-second windows. As compared to the rest data, a longer time was required for extracting the ensemble averages (20 seconds compared to 5 seconds) due to increased postural sway in exercise recovery compared to rest. An ensemble average, RI interval and PEP were calculated for each of the 15 windows. A correlation analysis was then performed between RI intervals and PEP among all 17 subjects.

2.2.7 Estimating Relative Changes in Stroke Volume

Stroke volume *SV1* was calculated for each subject during rest using the reference ICG signal using Sramek's equation [49]:

$$SV1 = \frac{(0.17H)^3}{4.25Z_0} \cdot \left(\frac{d_z}{dt_{\max}} \right) \cdot LVET \tag{1}$$

where H is the height of the subject, Z_0 is the base impedance of the thorax, dz/dt_{max} is the amplitude of the global maxima of the ICG (defined as the first derivative of the thoracic impedance as a function of time), and LVET is the left ventricular ejection time which is calculated as the timing difference between X-point and B-point in the ICG.

Then, for each of the 17 subjects separately, the 5 minutes after exercise recovery recording of head-to-foot force plate BCG and ICG were divided into 15 20-second ensemble averages (as with the PEP estimation described above). The normalized change in stroke volume ΔSV_i for each of the 15 ensemble averages was calculated as follows:

$$\Delta SV_i = \frac{SV_i - SV1}{SV1} \tag{2}$$

For each subject, subject specific correlations were performed between the J-peak of the force plate BCG and the X-point of the ICG and between the I-wave of the force plate BCG and the B-point of the ICG. Linear regression models were found for each subject and used to estimate the X-point, $\hat{X}$, and B-point, $\hat{B}$, in every ensemble average. Estimated LVET for each of the 15 ensemble averages for each subject was then calculated.

Similarly, another subject-specific correlation was performed between the amplitude of the J-peak and d_z/dt_{max}, and the linear regression model was used to estimate d_z/dt_{max} for the ensemble averages for that subject. Finally, the estimated ΔSV was calculated for each of the 15 ensemble averages for every subject and the resulting estimates were compared with the reference ΔSV calculated from ICG.

2.2.8 Signal-to-Noise Ratio Calculations

For each of the 17 subjects, an ensemble average $\bar{x}$ was obtained for the 60 second baseline recording of weighing scale BCG and head-to-foot force plate BCG. The 60 second window was then divided into 12 5-second sub-ensembles x_i and a sub-ensemble average $\bar{x}_i$ was obtained for each of the 12 sub-ensembles. 5-second ensembles were used rather than individual beats to reduce motion artifacts. Noise was calculated for each of the sub-ensemble averages using:

$$n_i = \bar{x}_i - a_i \bar{x} \tag{3}$$

where a_i is the normalization coefficient for the amplitude of the ensemble averaged BCG w.r.t the ith sub-ensemble average:

$$a_i = \frac{\bar{x}_i \cdot \bar{x}}{\bar{x} \cdot \bar{x}} \tag{4}$$

The noise-to-signal ratio (NSR) was then calculated for each of the sub-ensemble averages:

$$NSR_i = \frac{\mathrm{var}(n_i)}{\mathrm{var}(\bar{x})} \tag{5}$$

Finally, SNR was calculated as follows:

$$SNR = \frac{1}{\frac{1}{12}\sum_{i=1}^{12} NSR_i} \tag{6}$$

2.3 Results and Discussions

2.3.1 RI and PEP Correlation for Scale and Force Plate BCG during Rest

The correlation results were: $r^2 = 0.85$, $m = 0.97$, $b = 34.92$ for the force plate and $r^2 = 0.81$, $m = 0.92$, $b = 24.23$ for the scale, with *m* being the slope and *b* being the *y*-intercept. These results show an improved timing accuracy in force plate compared to scale BCG in estimating PEP, demonstrating that the limited mechanical bandwidth in the weighing scale impacts the timing accuracy. Figure 10(a) shows a linear regression fit for scale and force plate BCG RI-intervals vs PEP and Figure 10(b) is a Bland Altman plot [50] that shows a larger standard deviation in PEP estimated from scale BCG compared to PEP estimated from force plate BCG.

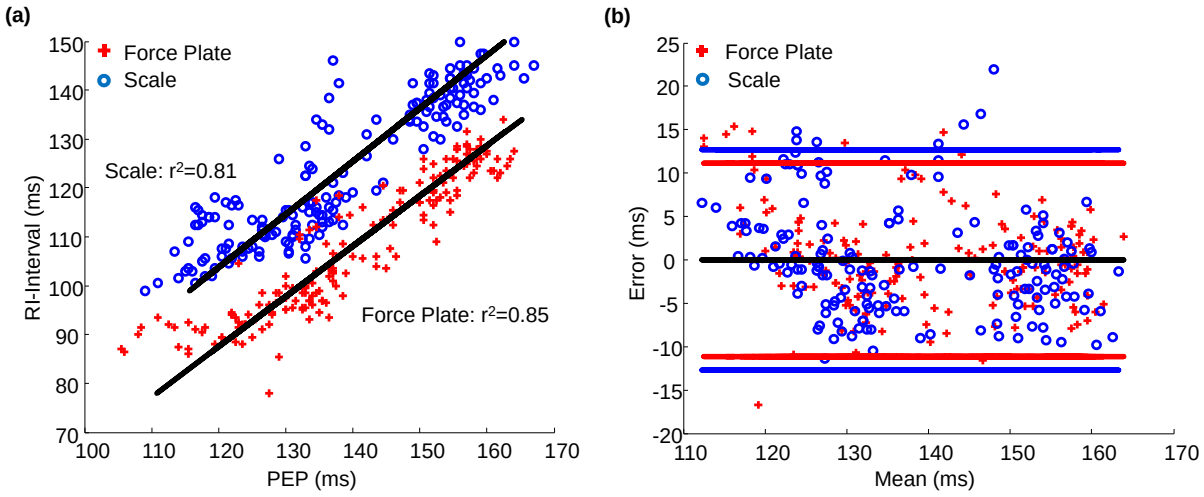

Figure 10. (a) Linear regression fit for both scale and force plate head-to-foot BCG RI-interval vs PEP among all subjects. (b) Bland Altman plot for scale and force plate linear prediction models of PEP. The blue line is the 95% confidence range of the PEP estimations from the scale BCG RI interval while the red line is the 95% confidence range of the PEP estimations from the force plate BCG RI interval.

2.3.2 RI and PEP Correlation for Head-to-Foot Force Plate BCG during Recovery

The results were: $r^2 = 0.92$, $m = 1.14$, $b = 14.46$. Figure 11(a) shows a linear regression fit for force plate head-to-foot BCG vs PEP during recovery. The higher correlation of head-to-foot BCG I-wave to PEP during exercise recovery compared to that during rest ($r^2 = 0.85$) can be attributed to the more significant changes in PEP and RI intervals during exercise recovery as both values increase to stabilize again at the baseline values. Additionally, higher frequency information in after exercise BCG, as can be observed in Figure 11(b), is another cause of increased accuracy over rest BCG. These results show a better correlation coefficient (0.92 versus 0.86) and a much smaller y-intercept (14.6 versus 138) than the results obtained from correlating RJ interval to PEP in the previous work [47]. One of the future work objectives in that paper was to examine waves other than the J-wave for better correlation and smaller y-intercept but that was not

feasible because of the difficulty of detecting those other waves during exercise recovery. However, with the force plate BCG, we were able to successfully extract the I-waves from all subjects and hence obtain the improved results.

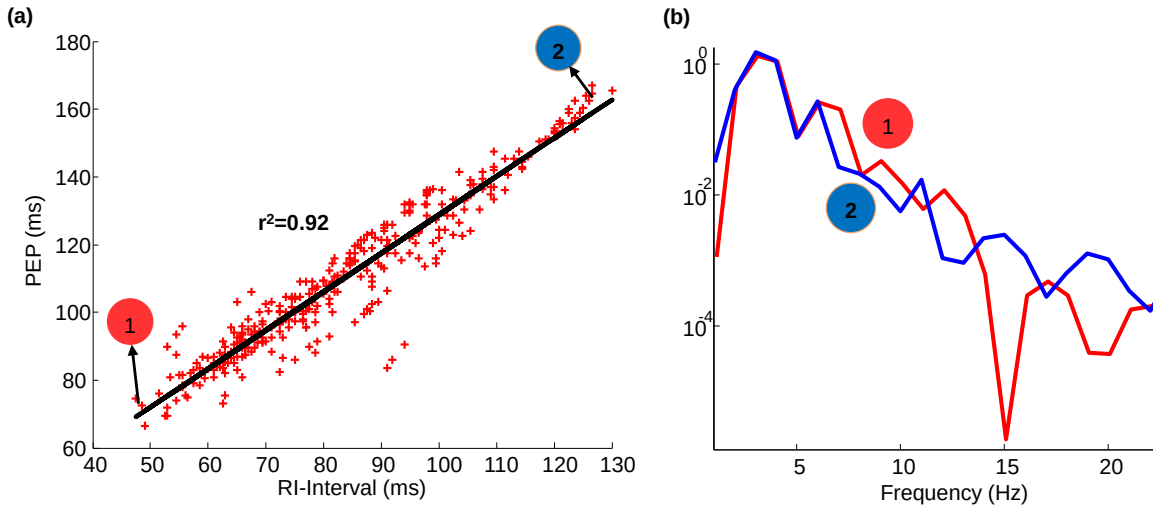

Figure 11. (a) Linear regression fit for head-to-foot force plate BCG vs PEP during recovery among all 17 subjects. (b) Power spectral density of the force plate BCG waveforms that correspond to the SV estimates marked in (a). Higher frequency information exists in after exercise BCG (waveform 1) compared to rest BCG (waveform 2).

2.3.3 Stroke Volume Estimation from Head-to-Foot Force Plate BCG during Recovery

The mean error among all subject was 5.3% with a standard deviation of 4.2%. We compared our results to other studies providing such detailed error analysis information on non-invasive SV estimation. In the ones available [51-54], the following results were achieved and considered "acceptable". In [51], the mean error was 16.5% for SV derived from femoral arterial blood pressure (ABP) and 14.5% for SV derived from radial ABP. In [53], the mean error was 8.9% for SV estimated using mean arterial pressure (MAP) and 14.35% for SV estimated using the systolic waveform. In [54], the percentage error between electric velocimetry and transesophageal Doppler echocardiograph for estimating

cardiac output (CO) –which is obtained from multiplying SV by heart rate – was 29%. All of these results were considered acceptable. In [52], the comparison between SV estimation using the Vigileo-FloTrac^{TM} system and esophageal Doppler yielded an error rate of 58% which was deemed unacceptable, since according to [55], limits of agreement of ±30% are acceptable. Hence, our method allows ΔSV estimations for every subject accurately and can provide insight in the case of HF patients.

Additional comparison with the cardiac output estimation methods that were summarized in [39] is shown in Table 1. It should be noted here that these works used different gold standards to compare their methods against. All non-invasive gold standards, including ICG, are susceptible to errors compared to invasive measurements such as thermodilution or Fick's method. However, we did not have access to the invasive measurements methods and ICG, which is often used in research, has proved good correlation with Fick's method. Therefore, when interpreting the results, we should take into consideration inherent errors in the gold standard. The per-subject results of the absolute and percentage mean and standard deviations of the estimated ΔSV compared to the reference calculations are shown in Table 2. Figure 12 also shows the percentage ΔSV in comparison with the calculated percentage ΔSV from ICG.

We can observe in Figure 12 that the change in SV is positive in some cases while negative in other cases. This can be attributed to the fact that SV increases during exercise and then decreases again during exercise recovery and it sometimes goes below the baseline SV by the end of the recovery period. Hence, at the beginning of the recovery period, the change in SV would be positive while at the end of the recovery period the change in SV would be around zero or sometimes negative.

Table 1. Comparison of several studies on non-invasive cardiac output measurement for exercise and exercise recovery (healthy subjects).

Author (year)	Method	Gold standard	# sub.	# data points	R^2
This work	**Ballistocardiography**	**ICG**	**17**	**255**	**0.96**
Inan et al (2009)	Ballistocardiography	Doppler echo	9	275	0.85
Houtman et al (1999)	Cont. Blood Pressure	Gas rebreathing	12	24	0.46
Sugawra et al (2003)	Cont. Blood Pressure	Doppler echo	16	640	0.76
Antonutto et al (1995)	Cont. Blood Pressure	Doppler echo	9	27	0.77
Christie et al (1987)	Doppler echo	Fick (direct)	10	42	0.66
Wilmore et al (1982)	Gas rebreathing	Thermodilution	6	12	0.76
Liu et al (1997)	Gas rebreathing	Fick (direct)	9	37	0.77
Johnson et al (2000)	Gas rebreathing	Fick (direct)	6	96	0.90
Zhang et al (1986)	ICG	Gas rebreathing	10	78	0.91
Miyamoto et al (1981)	ICG	Gas rebreathing	6	19	0.83
Moore et al (1992)	ICG	Gas rebreathing	11	44	0.76
Hatcher and srb (1986)	ICG	Gas rebreathing	60	230	0.56
Trodie et al (2004)	ICG	Gas rebreathing	8	40	0.82
Richard et al (2001)	ICG	Fick (direct)	12	50	0.88

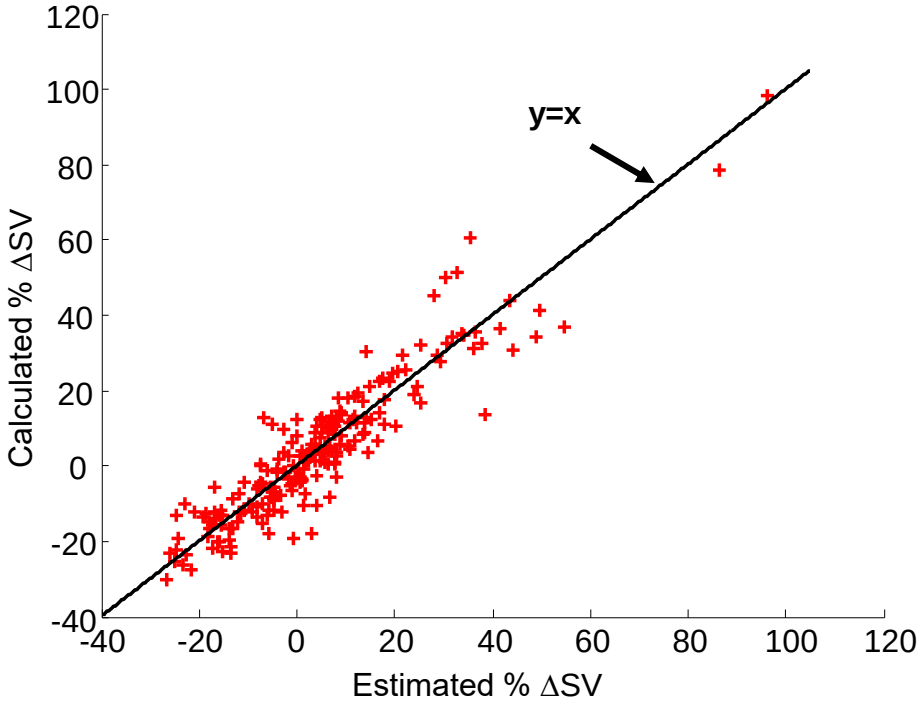

Figure 12. Estimated stroke volume percent changes from head-to-foot force plate BCG compared to calculated stroke volume percent changes from reference ICG.

Table 2. Per subject errors in ΔSV estimation.

Subject	μ Error (ml)	σ Error (ml)	% μ Error	% σ Error
1	2.0	1.7	3.6	3.1
2	1.4	0.9	3.4	2.3
3	3.7	2.6	6.5	4.6
4	4.0	4.9	5.4	6.7
5	2.5	1.8	6.4	4.7
6	2.8	2.0	8.9	6.4
7	3.4	3.1	6.6	6.0
8	2.8	2.6	4.6	4.3
9	5.6	4.9	11.0	9.5
10	3.4	2.5	4.5	3.4
11	3.1	3.0	5.3	5.1
12	3.3	1.4	5.2	2.1
13	1.6	1.3	3.3	2.8
14	1.9	1.4	4.6	3.3
15	1.0	0.8	2.2	1.8
16	3.4	1.8	4.7	2.5
17	3.5	2.5	4.6	3.2
Average	**2.9**	**2.3**	**5.3**	**4.2**

2.3.4 Signal-to-Noise Ratio Comparison for Scale BCG and Force Plate Head-to-Foot BCG

SNR results show slightly better SNR values for force plate head-to-foot BCG compared to scale BCG. The results for SNR calculations are shown in Table 3.

Table 3. Per subject SNR calculations for scale and force plate BCG.

subject	Gender	Height (cm)	Weight (kg)	FP SNR (dB)	Scale SNR (dB)
1	Female	160	59	6.0	5.6
2	Male	175	75	8.6	7.6
3	Female	168	68	8.1	5.3
4	Female	160	52	1.7	1.3
5	Male	183	86	8.3	4.3

6	Male	175	74	1.5	-1.8
7	Female	152	49	5.5	4.6
8	Male	178	65	3.4	1.7
9	Male	178	88	1.6	1.5
10	Male	178	68	2.6	1.2
11	Male	190	88	7.1	5.1
12	Female	175	68	9.0	8.1
13	Male	175	70	3.8	3.8
14	Male	185	76	8.3	8.3
15	Male	175	79	3.1	2.7
16	Female	163	75	6.7	6.6
17	Female	168	61	8.0	7.8

The force of the BCG signal is 2 Npp and the sensitivity of the transducer in the weighing scale is 19.1 μV/N [16] which would give a BCG voltage of 38.2 μVpp. The noise level referred to the input of the circuit is 150 nVpp. This would lead to an SNR of 255 in the linear scale or 48 dB. As for the force plate, with the same BCG amplitude of 2 Npp for the BCG and a sensitivity of 19 mV/N amplified by 250 times in the z-direction gives a signal of 9.5 Vpp. The output noise level is 5mVpp which leads to an SNR of 1900 in the linear scale or 66 dB. Those calculations are clearly far better from what we are obtaining in terms of SNR for both the scale and force plate so the difference cannot be attributed to this. Hence, it is likely that the BCG SNR is limited by motion artifacts, not electronic noise.

Moreover, we observed the order in which subjects stood on the scale and force plate (since half the subjects stood on scale first while the other half stood on the force plate first) to account for any bias due to the order in which the subject stood on the force plate versus the weighing scale, and found that the difference in SNRs between force plate and weighing scale BCG was bigger for all the subjects that stood on the force plate first (subjects: 2,3,5,6,8,10,11, and 12 in Table 3) compared to those who stood on the weighing

scale first (subjects: 1,4,7,9,13,14,15,16, and 17 in Table 3). For both groups, the results are statistically significant ($P<0.05$).

We hypothesized that the improvement in SNR in force plate BCG over scale BCG is a result of the subject being able to stand more comfortably on the force plate with their feet wider apart than with the scale. To further investigate this theory, we had 7 healthy subjects (5 males, 2 females; ages: 24.9±1.6 years; height: 174.6±12.7 cm; weight: 73.2±18.4 kg) stand on the force plate with feet 0, 7.6, 12.7 (same distance as for the scale), 22.9, 27.9 cm apart for 60 seconds each. In addition to BCG, we measured the ECG. Same filtering, R-peak detection, ensemble averaging, and SNR calculation methods were used as described earlier in this chapter.

Results showed that the ideal stance width from highest BCG SNR varied from subject to subject. four subjects had the highest SNR BCG with feet 22.9 cm apart, one with feet 27.9 cm apart, one with feet 12.7 cm apart, and one with feet 7.6 cm apart. No relation was found between the optimal distance and the height or weight of the subjects. However, on average, the highest SNR of 7.1 dB was found when the feet were 22.9 cm apart (wide) and the lowest SNR of 4.2 dB occurred when feet were together (i.e. 0 cm apart). The average results among all subjects are shown in Figure 13.

While these findings were not consistently better for distances wider than BCG scale stance for every subject, an important conclusion from this experiment is that in longitudinal studies, where BCG waveforms are compared from one day to the next, it would be important to control the distance between the feet. This could be accomplished, for example, by providing some reference indicators on the force plate platform for foot placement.

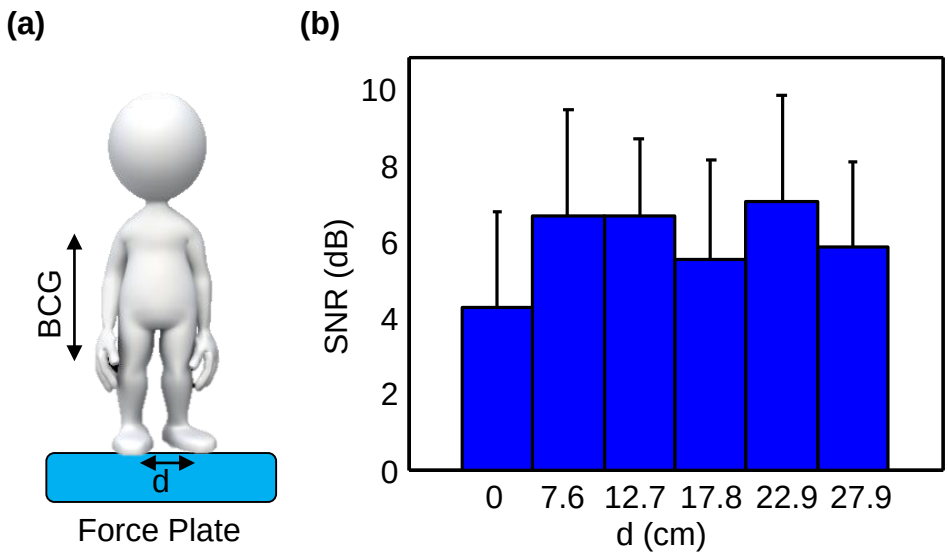

Figure 13. (a) a subject standing on the force plate based system for BCG signal measurement with feet d cm apart. (b) Average SNR across subjects for different stance widths. The lowest SNR was found with the feet placed together (0 cm), and the highest SNR was for a fairly wide stance (22.9 cm).

2.4 Conclusions

We have shown that force plate BCG allows for more reliable PEP estimations than scale BCG due to more accurate timing information that can be attributed to the wider bandwidth of the force plate. We have also shown that force plate head-to-foot BCG can be used to estimate ΔSV accurately. However, in order to determine subject-specific coefficients for how the BCG parameters relate to the ICG parameters, simultaneous ICG measurements are needed at first for calibration purposes. Hence, the estimation of ΔSV hinges on having the initial reference measurement. Nevertheless, for the HF patients monitoring application, such initial measurements of SV can be obtained at the index visit, and then the BCG-equipped force plate can be used to measure changes from that initial value over time. Changes in SV could then be used to titrate care, thus potentially leading to improved outcomes and reduced costs.

Nevertheless, the cost of the force plate is significantly higher than the weighing scale, and thus the force plate – at this current cost – would not be feasible for at-home use. If the force plate was positioned in a central location accessible to multiple HF patients such as in a grocery store, the cost of the force plate could be amortized over the large number of patients that could access the device, thus allowing it to be a feasible option for monitoring patients outside the clinic / hospital setting. A more feasible method of non-invasive and unobtrusive monitoring of HF patients would be to obtain SCG signals from wearable sensors. This would be more cost-effective and it would provide the health care providers with continuous PEP estimates from patients, which could potentially lead to a more efficient form of proactive care.

CHAPTER 3. CLASSIFICATION OF WEARABLE SEISMOCARDIOGRAM SENSOR POSITIONING FOR ROBUST ESTIMATION OF THE PRE-EJECTION PERIOD IN UNSUPERVISED SETTINGS

3.1 Introduction

While the modified weighing scale or force plate can potentially enable measurement of STIs outside of the clinical setting, measurements are limited only to when the subject stands on the platform. To facilitate more frequent measurements, and specifically measurements during periods of great interest clinically such as when a person is performing exercise (e.g., walking up a flight of stairs), we also investigated techniques for robust measurement of STIs using wearable devices. Importantly, it has been recently demonstrated that the PEP measured using ECG and SCG waveforms taken by a small wearable patch before and after walking can be used to distinguish between compensated and decompensated HF patients, indicating that the measured SCG signal quality is sufficiently high for this clinical population for cardiovascular health assessment purposes [46].

However, one key challenge with using SCG signals for PEP estimation is that the position of the accelerometer on the chest may affect the relationship between the "AO" point detected on the SCG signal and the ground truth aortic valve opening timing detected by a reference standard measurement (e.g., an ICG). For mid-sternal SCG measurements, the AO point corresponds to the B-point of the ICG closely in time across different subjects and thus can directly be used as the endpoint for PEP time interval estimation (with the ECG Q-wave being the starting point). However, since SCG is a measure of local accelerations of the chest wall, changing the positioning of the sensor can change the shape

of the SCG signal [56], and thus the relationship between the AO point detected by the SCG and the actual timing of aortic valve opening. Specifically, the regression curve linking Q-AO to Q-B intervals across subjects would have different slope and y-intercept for each position on the chest for SCG measurements. Thus, if the position of the accelerometer on the chest is not known a-priori, the timing intervals extracted from the SCG signals may not accurately reflect the underlying physiology of the heart.

This positioning problem is not an issue when using BCG since it is a measure of entire body vibrations rather than local chest vibrations. However, the main advantage of using SCG over BCG is the possibility of continuous monitoring, which can allow for monitoring in the context of normal stressors encountered by the person during their everyday activities (e.g., analyzing the changes in the SCG in response to exercise or other stressors as opposed to simply a once-per-day measurement by the scale).

It is essential to study how research efforts cross domains from clinical to non-clinical settings, and vice versa, to verify that both are congruent [57]. Therefore, we quantitatively study how the positioning of the accelerometer affects the slope and y-intercept of this regression function (linking Q-AO to Q-B across subjects) and devise a method that enables the person using the wearable device to know whether or not the accelerometer is placed in the ideal (mid-sternal) position. Hence, the user will be able to adjust the positioning of the device and avoid computation of inaccurate estimates for the days on which the device is misplaced. This will be an important step in allowing ECG and SCG measurements to be used, for example, by HF patients at home for monitoring changes in the underlying condition and adjusting therapies accordingly.

3.2 Methods

3.2.1 Protocol

The study was conducted under a protocol reviewed and approved by the Georgia Institute of Technology Institutional Review Board (IRB). All subjects provided written consent before experimentation. Data were collected from ten healthy subjects: five females and five males (demographics: 24.7±2.3 years, 170±11.6 cm, 70±10.5 kg). The experiment was performed on two consecutive days. On the first day, each subject was asked to stand still for 60 seconds with accelerometers placed on the mid-sternum, approximately 7.5 cm to the right, and 7.5 cm to the left to obtain baseline measurements. The subject then performed a stepping exercise for one minute after which he was asked to stand still for 5 minutes to monitor the recovery. On the second day, the same protocol was repeated except the accelerometers were placed on the mid-sternum, 5 cm above, and 5 cm below (selected to be slightly less distance than the horizontal "misplacement" to closer simulate expected conditions in practice). On both days, ECG and ICG signals were collected as well. Figure 14(a) shows the sensor placement on both days combined and Figure 14(b) shows the SCG signals obtained from the different placements of accelerometers.

3.2.2 Hardware Design

The ECG and the ICG signals were measured using the same hardware described in section 3.1.2. To detect the small body vibrations in response to cardiac ejection (i.e., SCG signals), we used ultra-low noise and small-footprint 356A32 instrumentation accelerometers. The accelerometers are powered by the Model 482C15 sensor signal

conditioner (PCB, Piezotronics, Inc., Depew, NY). Only the signals along the dorsoventral direction from the accelerometers were analyzed in this study.

3.2.3 Data Processing

The SCG, ICG and ECG signals were filtered with an FIR Kaiser window band-pass filters (cut-off frequencies: 0.8-30 Hz for SCG, 0.8-35 Hz for the ICG and ECG). The same method described in section 2.2.3 was used for ensemble averaging.

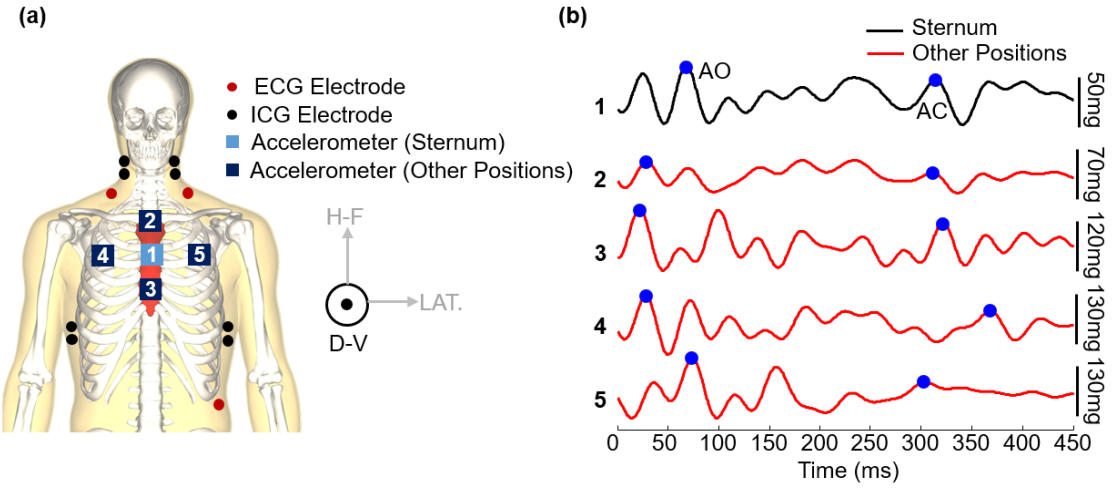

Figure 14. a) Adapted from [58]. An illustration of the experimental setup. (b) A 450 ms portion of one ensemble averaged beat of an SCG signal collected from the accelerometer placed on the mid-sternal position vs SCG signals collected from the other accelerometers at rest. It can be observed that the waveforms are very different in their shapes and result in different estimations of the aortic valve opening (AO) and aortic valve closure (AC) points.

3.2.4 Feature Extraction

There is a lack of well-defined standards for the fiducial points of cardiomechanical signals, including SCG, due to how the morphology of the signal is affected by age, sex, and heart conditions. In a recent study, a delineation algorithm was designed to detect the fiducial points of the SCG signal, including the AO-point, by training an algorithm on

48,318 manually annotated cardiac cycles and resulted in good accuracy when tested on healthy individuals [59]. Additionally, many of the papers in literature detected the AO-point as the second positive peak of the SCG signal [33], However, we have found with our measurements, with standing rather than supine subjects, that this peak is not necessarily reproducible from person to person, and the noise in the measurements can easily corrupt this peak's detection as compared to the detection of the largest peak (positive or negative). Hence, to detect AO from the dorsoventral SCG signals, the timing of the minimum or maximum absolute magnitude in the first 200 ms is detected [60]. The B-point is detected as described in section 3.1.4.

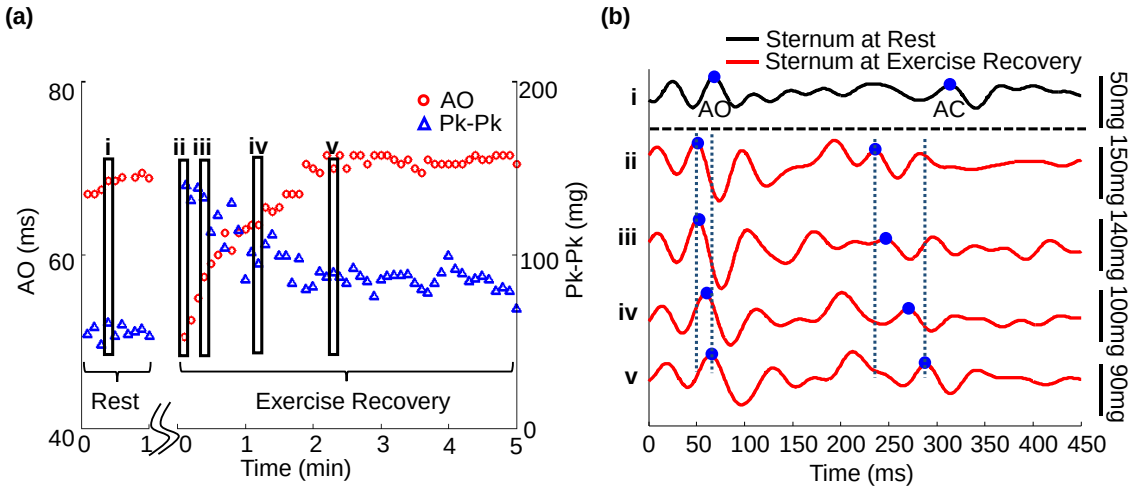

Figure 15. (a) The variation in AO and peak-to-peak amplitude of the DV SCG signal collected from mid-sternum as the physiological state of the subject changes between rest and exercise recovery. It can be noted that right after exercise the AO time decreases while the peak-to-peak amplitude increases only to stabilize back to baseline value as time passes. The five rectangles represent the AO and peak-to-peak values of the five waveforms shown in part (d) for the different physiological states. (b) A 450 ms portion of an ensemble average of DV SCG beats collected from the sternum during rest and different states during exercise recovery. We can observe that the waveforms look similar for a time shift and an increase in amplitude after exercise.

Figure 15(a) shows how the AO and Peak-to-Peak values vary between the rest and exercise recovery physiological states and Figure 15(b) shows example SCG waveforms

from one accelerometer location at different physiological states. The first waveform in Figure 15(b) is an SCG waveform taken when the subject is at rest while the rest of the waveforms are taken at different times during the post exercise recovery. The waveforms are identical in shape and the only change is a time shift to the left (i.e, AO and AC points occur at earlier times in the heart beat cycle) and an increase in the peak-to-peak amplitude of the waveform. By the end of exercise recovery, the SCG signal becomes identical to the original baseline rest SCG signal.

3.2.5 Correlation During Rest and Exercise Recovery

To quantitatively investigate how the change in the AO point of SCG for different accelerometer positions affects PEP estimation, a correlation analysis was performed between R-AO intervals and R-B intervals for the different sensor positions. Every 6 beats of the signals at rest were averaged together to obtain ensemble averages with 50% overlap and every 10 beats of the signals during exercise recovery were also averaged together to obtain ensemble averages with 50% overlap. Six beats were sufficient for noise reduction when the subject was at rest, and was long enough to capture a typical respiratory cycle, while a higher number of beats was used for ensemble averaging during exercise recovery because of the increased postural sway as compared to rest. PEP (the timing at which the B-point – the opening of the aortic valve – occurs) was calculated. Additionally, the R-AO intervals of all SCG ensemble averages were calculated. Linear, quadratic and cubic regression were then performed with the y-value being the R-B interval (PEP) and the x-value being the R-AO interval and the correlation coefficient *r*, was found for each type of regression.

3.2.6 Classification of Different Positions of Accelerometers

To automatically determine when an accelerometer is placed in the desired, mid-sternal, position or a different position, we implemented binary classification. This is done by extracting features from the SCG waveforms of all positions, labeling instances obtained from the mid-sternal SCG differently than instances obtained from SCG from other positions, dividing the instances we have into training and testing sets, and training a classifier on the training sets in order to predict the class of the instances in the test sets.

The same ensemble averages of the SCG signals used for the AO point detection were used for extracting features from these signals; specifically, a set of 26 temporal and spectral features were extracted. Those features consisted of: mean, kurtosis, skewness, median, standard deviation, peak-to-peak (0-250 ms and 250-500 ms), RMS and peak to RMS ratio (0-250 ms and 250-500 ms) as temporal features and band power in the following bands (0-3 Hz, 3-6 Hz, 6-9 Hz, 9-12 Hz, 12-15 Hz, 15-18 Hz), 1st, 2nd and 3rd peak of the PSD, frequency of the 1st, 2nd and 3rd peak of PSD, mean of the PSD, kurtosis of the PSD, skewness of the PSD, and standard deviation of the PSD as frequency domain features. Table 4 shows the top features in terms of the average information gain provided by each feature, with information gain being a measure of the reduction in entropy (a measure of impurity) of the class variable after the value for the feature is observed, i.e.:

$$IG(T,a) = H(T) - H(T|a) \tag{7}$$

where H is the entropy function, T are the training instances and α is the attribute (feature) in question [61]. The entropy function $H(X)$, which is the expected number of bits needed to encode a randomly drawn value of X, is a measure of impurity of the set of training examples and is given by:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \quad (8)$$

Table 4. Best features in terms of information gain

Info Gain	Feature	Feature Type
0.053±0.014	Peak-to-Peak (1st 250 ms)	Time
0.049±0.011	Standard Deviation	Time
0.049±0.011	RMS	Time
0.046±0.007	1st Peak of the PSD	Frequency
0.044±0.003	Mean of the PSD	Frequency
0.043±0.006	Peak-to-Peak (2nd 250 ms)	Time
0.043±0.005	Standard Deviation of the PSD	Frequency

Figure 16 is an illustration of some of those features in the time and frequency domain.

The above mentioned features were extracted from SCG ensemble averages for all accelerometers' positions for each subject. Instances from SCG acquired with an accelerometer in the mid-sternal position were associated with the label -1 while instances from all other positions were associated with the label +1. Multiple training and testing sets were created. For every subject to be tested, the training set included rest and exercise recovery instances from all other subjects, half the rest and exercise recovery instances from the SCG acquired with the accelerometer placed on the mid-sternum for that subject, and only rest instances from SCG acquired by accelerometers from all other positions for the tested subject. Four separate test sets were created with each of them consisting of the other half of the rest and exercise recovery instances of the mid-sternal SCG and the exercise recovery instances from SCG acquired by one of the differently placed accelerometers. This division resulted in a total of 10 training sets (one per subject) and 40 test sets (4 per subject for each of the 10 subjects).

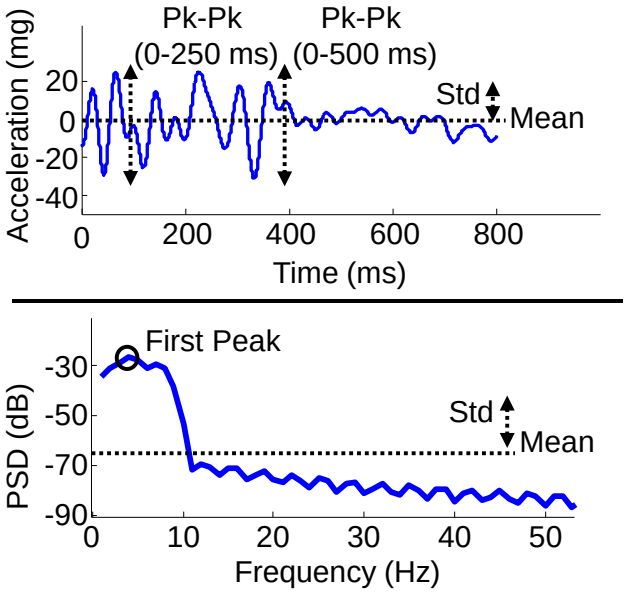

Figure 16. An illustration of some of the top features in the time and frequency domain.

We exclude the exercise recovery instances from all positions different than the mid-sternal position from the training sets for the subject being tested, because we assume that we will not have that information when the device is initially calibrated for a particular subject. More specifically, when the subject wears the device for the first time under the supervision of a medical professional, it will be placed in the ideal position and several other positions for a duration of 60 seconds each. The subject (who may, for example, be a patient with HF at the index visit) will then be asked to perform the six-minute walk test [62] with the device placed in the mid-sternal position (which justifies using some of the exercise recovery data from the mid-sternal position). These initial measurements that are collected only when the patient uses the device for the first time, are used for calibration purposes, and hence, can be included in the training set. An illustration of the division of features into training and testing sets for one subject is shown in Figure 17.

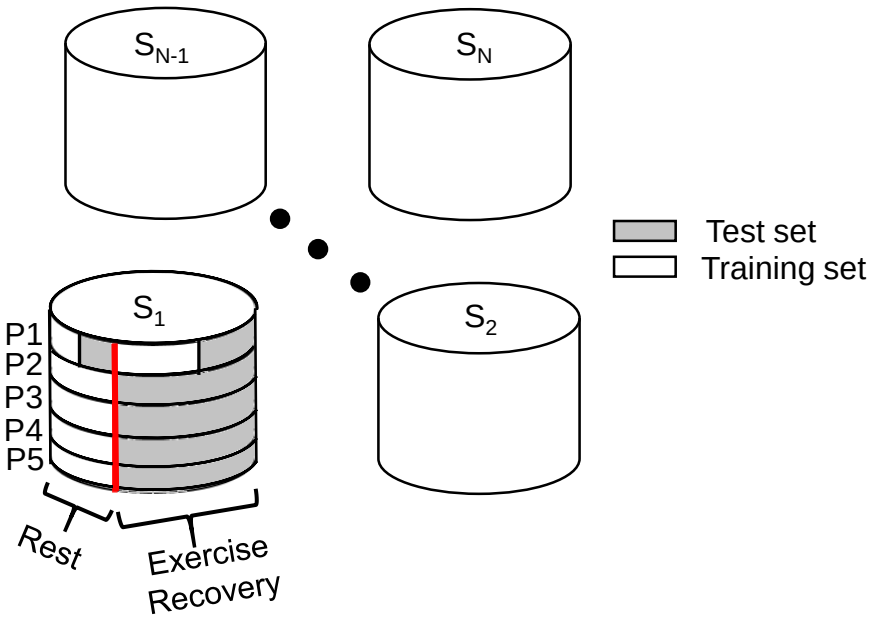

Figure 17. The training and testing sets for training and testing the classifier on one of the subjects. The shaded areas are features that are excluded from the training set and included in the testing sets when trying to detect the sensor position for the subject.

A boosted J-48 decision tree classifier with Adaptive Boosting algorithm was trained on each of the training sets and tested on the corresponding test sets for each subject to predict whether the SCG features are acquired from an accelerometer placed in the mid-sternal position or other positions [63, 64]. There are many advantages for using the J-48 Decision tree classifier as compared to other classifiers: 1) Decision trees are not sensitive to outliers since pruning is applied [61]; 2) Decision trees are non-parametric thus alleviating concerns about the data being linearly separable; 3) There is no need to tune different parameters as is the case with support vector machines, for example, which enables implementation of the same classifier for every subject; 4) Decision trees implicitly implement a form of feature selection by using the features with higher info gain as the top nodes in the tree [61]. Moreover, adaptive boosting improves the performance of the

decision tree by running the classifier a number of iterations (10 in our case) and improving it by accounting for incorrectly classified instances from the training set at each iteration.

3.3 Results and Discussions

3.3.1 PEP and R-AO Correlation Results for SCG from Different Sensor Positions

For each subject, the correlation coefficient resulting from the linear, quadratic and cubic regression between R-AO and PEP was calculated for every position. Results show high correlation for all positions which indicates that all positions are usable. However, each position, for the same subject, has its own unique regression curve relating to the B-point of the ICG, and we need to have the initial regression parameters for each position and be able to detect exactly where the sensor is placed each day, which is not the case, in order to be able to use all the positions. Thus, if a person wears the device on successive days, and positions the device in different places on different days, even with the same underlying R-wave to aortic valve opening timing, the measurements from the SCG itself may be completely different and one would not know if the person's underlying cardiovascular function changed, or if the device position changed, in comparing data from one day against the other.

This demonstrates that a position detection algorithm is needed to distinguish whether the wearable device is placed on the mid-sternum or not. Specifically, using the initially calibrated parameters, obtained from the linear, quadratic, and cubic regression functions used with the mid-sternal SCG to estimate PEP, when the sensor is placed somewhere else, would result in inaccurate estimations.

A summary of the obtained correlation coefficients, r, averaged across all subjects is shown in Table 5 and the regression curves for one of the subjects when the accelerometer is placed on the mid-sternal position is shown in Figure 18. Additionally,

Table 6 shows the average of absolute PEP estimation errors (ms) resulting from using the initially calibrated parameters obtained from the mid-sternal position when the accelerometer is placed in each of the 5 positions.

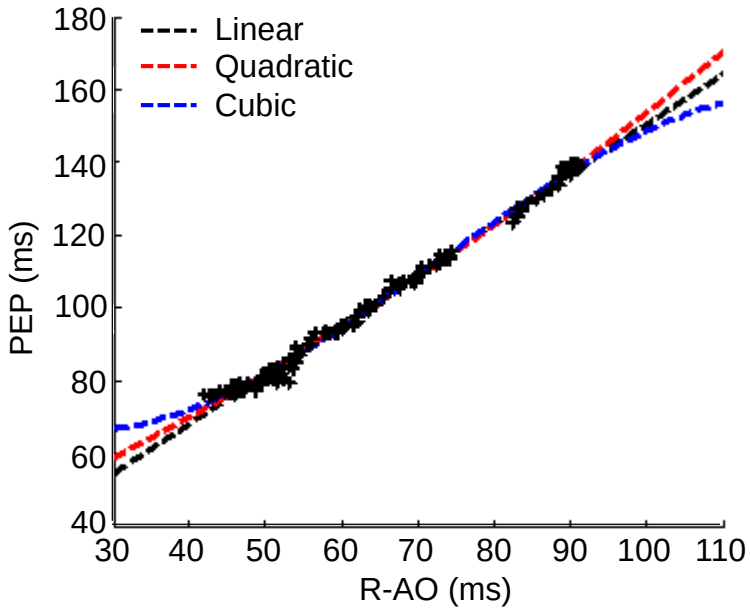

Figure 18. Linear, quadratic and cubic regression curves for one subject with the accelerometer placed on the mid-sternal position

Table 5. The average correlation coefficient (r) obtained from linear, quadratic, and cubic regression for the five different positions across all subjects.

	Correlation Coefficient (r)		
	Linear	Quadratic	Cubic
P1	0.96±0.027	0.96±0.042	0.97±0.038
P2	0.94±0.084	0.98±0.024	0.99±0.006
P3	0.95±0.048	0.96±0.049	0.96±0.049
P4	0.96±0.032	0.96±0.042	0.97±0.035
P5	0.94±0.078	0.95±0.072	0.95±0.067

Table 6. Average absolute PEP estimation errors (ms) for each position

		Subject										Avg.
		1	2	3	4	5	6	7	8	9	10	
Linear Regression	P1	2.5	3.4	5.8	1.2	5.0	2.7	2.9	2.0	5.9	4.2	3.6
	P2	4.8	4.7	16.7	5.3	52.9	10.9	27.9	10.1	167.1	6.3	30.7
	P3	22.6	6.5	8.7	11.8	79.8	10.7	105.4	35.9	107.1	5.8	39.4
	P4	24.9	10.4	60.4	18.1	46.3	15.2	28.5	54.9	5.2	21.3	28.4
	P5	8.7	12.1	252.6	6.7	41.2	12.9	54.8	84.4	10.8	6.1	49.0
Quadratic Regression	P1	1.5	1.7	5.5	1.1	5	2.6	2.7	1.8	3.4	3.3	2.9
	P2	5.0	5.8	19	2.9	52.5	10.6	30.6	9.7	106.4	9.5	25.2
	P3	21.3	5.6	19.2	5.7	79.0	10.4	130.5	29.6	115.2	11.1	42.8
	P4	27.4	11.9	31.6	16.2	125.8	14.4	28.4	33.9	8.3	30.4	32.8
	P5	9.6	12.5	807.8	6.4	109.1	11.6	54.6	35.7	12.8	6.0	106.6
Cubic Regression	P1	1.4	1.6	5.2	0.9	4.2	2.4	2.5	1.7	1.8	3.3	2.5
	P2	5.1	5.9	16.2	3	345.8	11.3	20	6.5	150.1	8.4	57.2
	P3	38.0	6.1	18.9	5.6	830.3	10.6	222.8	46.2	143.7	10.5	133.3
	P4	24.7	11.7	472.6	15.2	78.1	14.0	17.9	101.1	7.5	16.8	76
	P5	8.1	12.8	22358.0	4.5	137.4	11.6	33.8	260.0	12.6	5.6	2284.4

3.3.2 Classification Results

To quantify the performance of the classifier, we used precision and recall as our metrics. Precision (or positive predictive value) is defined as the fraction of retrieved instances that are relevant while recall (or sensitivity) is defined as the fraction of relevant instances that are retrieved [65]. In other words, a recall of 1, would indicate that we are detecting every time the sensor is placed in any position other than the mid-sternal position, and a precision of 1 would indicate that all of those detected instances are actually in a position different than the mid-sternal one. Equations (9) and (10) are the mathematical formulas for precision and recall respectively:

$$P = \frac{tp}{tp + fp} \tag{9}$$

$$R = \frac{tp}{tp + fn} \tag{10}$$

where *tp* is the number of true positives (instances correctly classified as belonging to the undesired class), *fp* is the number of false positives (instances incorrectly classified as belonging to the undesired class), and *fn* is the number of false negatives (instances incorrectly classified as belonging to the desired class).

With our classifier, we obtained a precision of 0.83 and a recall of 0.82. Considering that the classifier should return a +1 when a misplacement is detected and a -1 when no misplacement is detected, then, if the user is wearing the device every day of the year, i.e. 365 days, and is placing it in a position other than the desired one 100 days, we will be able to detect the incorrect placement and notify the user on 82 of these days. On 17 of the remaining 265 days, the user will be incorrectly notified to check the position of the device. This inconvenience of having the patient double check the position of the device on 17 days is insignificant when compared to the additional 82 days of correct PEP estimations that would have otherwise been inaccurate and could have potentially resulted in incorrect medical decisions.

It is worth mentioning that gender based anatomical differences may impact the results for female subjects compared to male subjects. Hence, the training and testing of the classifier was performed again after excluding all instances related to positions 4 and 5 in Figure 14(a), since they are the positions that are most likely to cause differences in signals based on gender and are also less likely to occur than positions 2 and 3. The results obtained were a precision of 0.83 and a recall of 0.81, which indicates that the inclusion or exclusion of those positions does not affect the results by much.

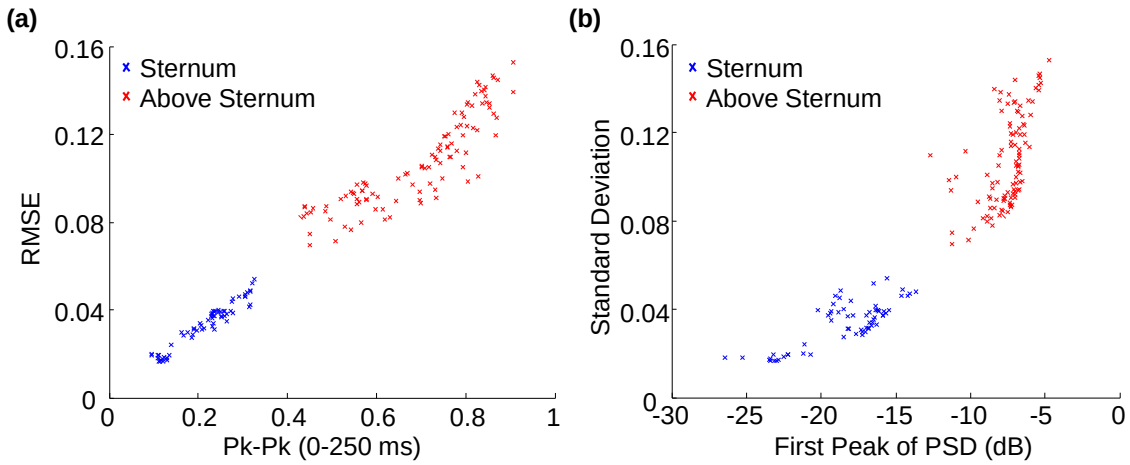

Figure 18. (a) The best feature in terms of info gain (Peak-to-Peak in the 1st 250 ms) plotted against the 3rd feature (RMS) for above sternum SCG compared to mid-sternum SCG in one of the test sets (b) The 2nd feature (Standard deviation) plotted against the 4th feature (1st peak of the PSD) for above sternum SCG compared to mid-sternum SCG in one of the test sets. It is clear in both (a) and (b) that those features result in two distinct clusters for these two positions which indicates that the position of the sensor (accelerometer) can be accurately detected.

From Table 6, we can observe that for subject 4, for example, calculations of PEP detection error resulted in an absolute average error of 1.2 ms when the device is correctly placed, 18.1 ms when it is placed to the right of the sternum, 6.7 ms when it is placed to the left of the sternum, 5.3 ms when it is placed in an upper-sternum position, and 11.8 ms when it is placed in a lower-sternum position. Each of these errors can be reduced to an average error of 1.2 ms in PEP estimation every time a wrong placement of the wearable device is detected. One should consider the possibility that day-by-day measurements would be compared against one another, but that the subject would position the device differently from one day to the next. In this case, having a different placement would lead to the R-AO interval meaning something very different on day 2 compared to day 1, etc.

It is also worth noting that the peaks that characterize the SCG waveform shape obtained from the mid-sternal position change at other positions, rendering it difficult to

even identify the main peaks of the waveform in certain positions. Therefore, detection of SCG sensor misplacement could also be useful for improving the robustness of SCG measurements in unsupervised settings overall, for parameters beyond PEP.

Figure 18(a) and Figure 18(b) show a clear separation of features from instances from a mid-sternal SCG compared to those of an above sternum SCG in the form of two distinct clusters.

3.4 Conclusions

We have shown that when SCG is used for PEP estimation, placing the accelerometer on different positions on the chest area results in different regression parameters (different slopes and intercepts). Hence, using the already calculated regression parameters from an SCG measured from the ideal position (mid-sternal position), for a certain subject, would result in inaccurate PEP estimates if the sensor is placed in a slightly different position. Therefore, we devised a method to detect when the sensor is placed in a position different than mid-sternum, in order to be able to provide feedback to the user to re-position the wearable device and avoid inaccurate measurements (or use the information regarding incorrect placement in post-processing to analyze the data accordingly).

In a practical scenario, when the wearable device is used by an HF patient, an initial calibration will be conducted first under the supervision of a medical professional, to calculate the regression parameters and obtain the training instances. After that, the patient will be wearing this device every day, which would enable the physicians to continuously monitor the patient's PEP, and as a result, the contractility of his heart. Additionally, the patient will be warned when the algorithm detects a sub-optimal placement of the device and will be asked to reposition it to guarantee accurate estimation of PEP. Such estimations

of PEP at home, in particular in response to activity, can potentially allow early detection of decompensation, and the changing of therapies to avoid hospitalization.

This chapter solves the issue of detecting the misplacement of the SCG sensor, and hence, avoiding inaccurate PEP estimation and wasted measurements. However, this does not eliminate the cause of the problem, which is that the morphology of the SCG signals changes with different placements, therefore changing the AO-point that is used to estimate PEP. If PEP can be estimated without relying solely on AO-point detection, but rather on multiple features extracted from the signal, then these estimates would be more robust to positioning changes. This is the motivation behind the work presented in the next chapter.

CHAPTER 4. UNIVERSAL PRE-EJECTION PERIOD ESTIMATION USING SEISMOCARDIOGRAPHY: QUANTIFYING THE EFFECTS OF SENSOR PLACEMENT AND REGRESSION ALGORITHMS

4.1 Introduction

An important limitation of SCG signals remains: SCG waveforms vary significantly from person to person which makes it difficult to accurately detect the "AO" point, and thus extract an absolute measure of PEP (i.e., the R-AO interval). Figure 19 shows three example SCG waveforms from different subjects with simultaneously obtained ICG signals. While the standard approach based on the existing literature is to denote the AO point as the second main peak of the SCG signal following the R-wave of ECG, this point does not occur concurrently with the "reference-standard" aortic valve opening point from the ICG signal (the B-point). For patients with HF, even greater variability is observed in the shape of the SCG waveform from person to person, thus motivating the need for improved methods for universal PEP estimation from SCG waveforms.

In this chapter, we set forth a novel approach to detecting the AO point from SCG recordings. Rather than searching for a single peak in the signal that corresponds to the aortic valve opening, we extracted multiple timing features from the SCG signal and combined all the detected features to create a universal regression model that predicts the relationship with the B-point in the ICG, across all subjects. We also quantified the effects of sensor location on the quality of PEP estimation to provide data-driven recommendations on optimal placement for minimizing PEP estimation error.

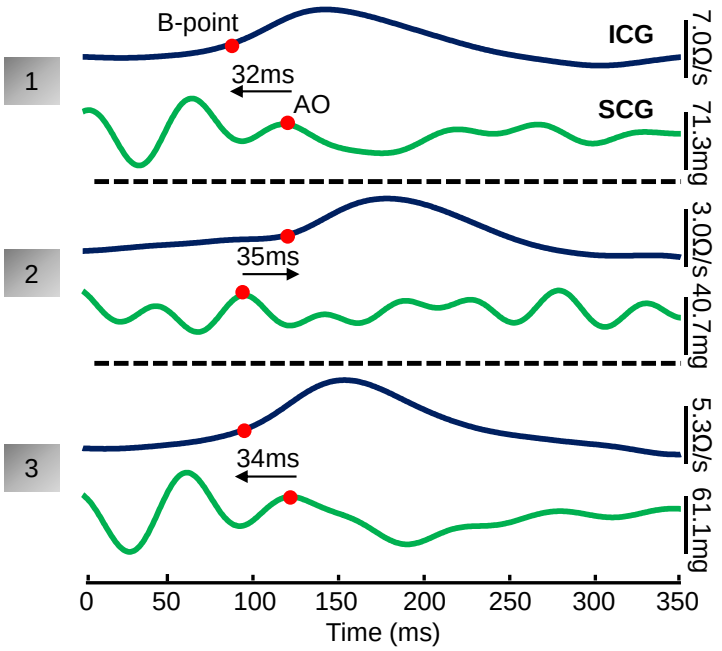

Figure 19. ICG and dorsoventral SCG ensemble averaged traces (n = 5 heartbeats) obtained with the sensor on the sternum for three different subjects. The ICG B-points and SCG AO-points are marked with red circles, and there is a substantial time difference between the two corresponding points for the three subjects: in two cases, the ICG B-point occurs first, and in the third case the SCG AO-point occurs first.

4.2 Methods

4.2.1 Protocol

The study was conducted under a protocol reviewed and approved by the Georgia Institute of Technology Institutional Review Board (IRB). All subjects provided written consent before experimentation. Data were collected from 10 healthy subjects, 5 males and 5 females (demographics: 23±3.3 years, 168.1±10.4 cm, 64.1±11.8 kg).

The experiment was divided into three parts with 15-minute breaks in between the parts. In each part of the experiment, the subject was asked to stand still on the BCG scale for one minute, then perform a stepping exercise for one minute, then stand still on the scale for five minutes (recovery). In part I, four accelerometers were placed on the subject:

one on each of the mid-sternum (Str), point of maximum impulse (PMI), below the left clavicle (LC) and below the right clavicle (RC). In part II, two accelerometers were coupled with a rigid plastic mold that was placed on the sternum such that the sensors were on the upper sternum (US) and lower sternum (LS). Finally, in part III, one accelerometer was placed on a flexible silicone sheet which was placed on the sternum such that the accelerometer was on the mid-sternum (FStr). In addition to BCG and SCG, we also measured electrocardiogram (ECG) signals, whose R-peaks were used as timing references for beat segmentation, and ICG signals, whose B-points were used as the reference standard timing for aortic valve opening. Figure 20(a) shows the placement of the electrodes and accelerometers in part I of the experiment. The purpose of part I of the experiment was to determine the location or combination of locations for SCG that provide the best PEP estimates. Parts II and III in combination with the sternum-placed accelerometer from part I are used to determine the best interfacing method of the sensors with the body.

4.2.2 Hardware and Data processing

To measure SCG signals, ADXL354 low noise, low drift, low power 3-axis accelerometers were used. The accelerometers were placed in a plastic case 2.8 cm wide and 3 cm long (shown in Figure 20(a)). The outputs of the accelerometers were connected to the data acquisition system (MP150, BIOPAC System, Inc. Goleta, CA). For part II of the experiment, two of the ADXL354 accelerometers in the plastic casing were placed in custom-made rigid acrylonitrile butadiene styrene (ABS) plastic mold that is 12.7 cm long with a 5.1 cm spacing between the accelerometers. For part III of the experiment, one of

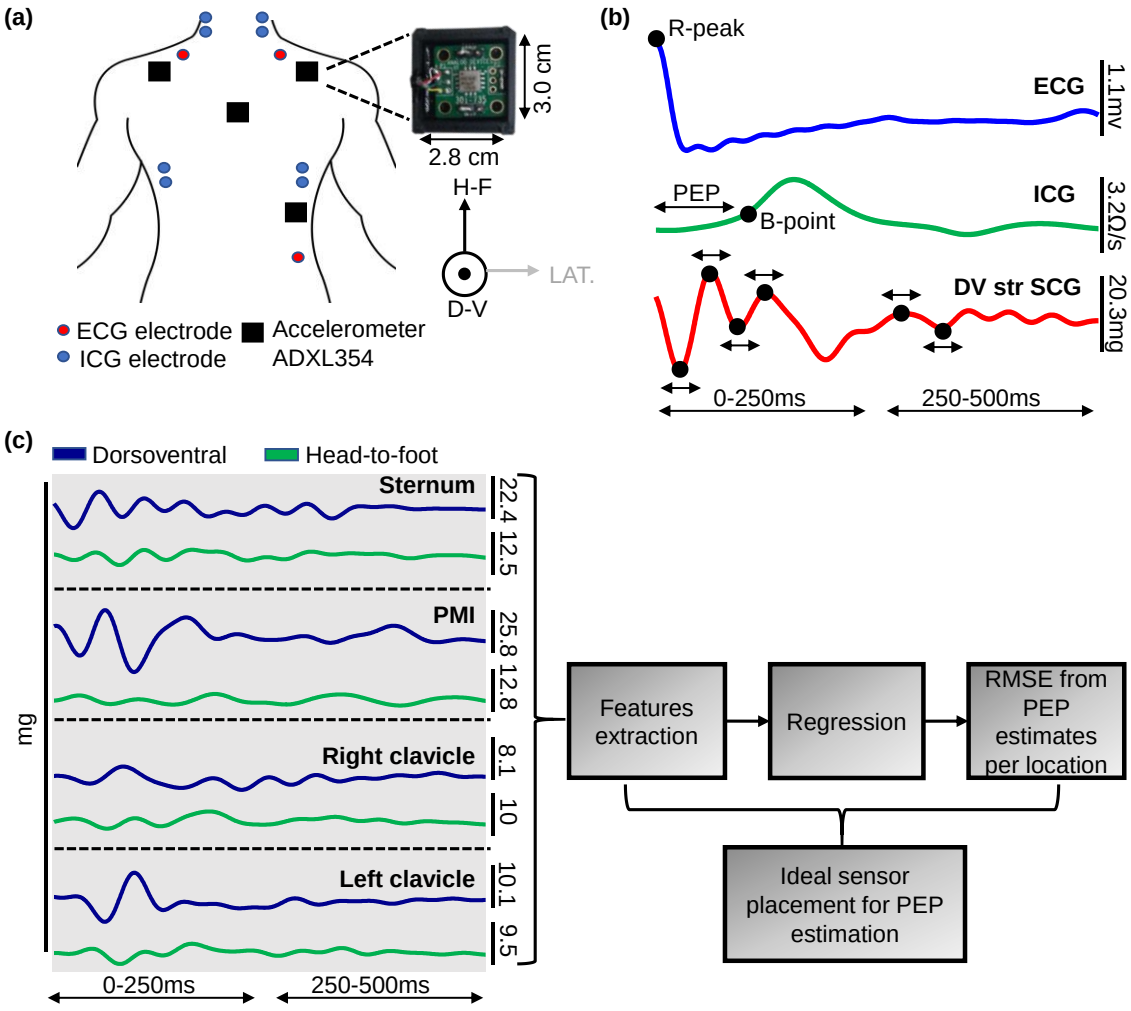

Figure 20. (a) The experimental setup for Part I of the experiment. Four ADXL354 Accelerometers are placed on the subject, one each at the mid-sternum, below the left and right clavicle, and at the point of maximal impulse. ICG and ECG signals are collected simultaneously. (b) Five beat ensemble averaged traces of ECG, ICG and mid-sternum dorsoventral SCG heartbeats. The ECG R-peak is used as a reference point for beat segmentation, the B-point of the ICG is used to detect aortic valve opening and the R-B interval is used as the ground truth PEP. Peak timing locations and width are extracted from the SCG signal as shown. (c) After extracting the features from the head-to-foot and dorsoventral axes of the SCG signals from all locations, a regression model is used to obtain PEP estimates from the features obtained from a single location, multiple combination of locations, one axis, and both axes. RMSE between the ground truth PEP and every estimate is calculated and the optimal location / combination of location and axes is determined.

the accelerometers in the plastic casing was placed on a silicone rubber sheet with durometer of 50A (Shore Hardness Scale) and dimensions of 15.2 cm length, 5.1 cm width

and 0.16 cm thickness. The different interfaces are shown in Figure 21(a).

To measure ECG and ICG, BN-EL50 and BN-NICO wireless modules (BIOPAC System, Inc. Goleta, CA) were used and to measure BCG the modified electronic weighing scale (BC534, Tanita, Tokyo, Japan) that was developed in [47] was used. All signals were sampled at 2 kHz.

The SCG, ICG, ECG and BCG signals were filtered with an FIR Kaiser window band-pass filters (cut-off frequencies: 0.8-30 Hz for the SCG, 0.8-35 Hz for the ICG, 0.8-40 for the ECG, and 0.5-20 Hz for the BCG).

4.2.3 Feature Extraction

Heart beats segmentation was performed as described in section 2.2.3. Every five beats from each of the SCG, BCG, and ICG signals were averaged together to reduce the noise, and an overlap of four beats between consecutive ensemble averages was used to maximize the number of ensembles. This made up a total of 6084 ensemble averages across all 10 subjects.

The B-point (inflection point) of the ICG was extracted as described in section 2.2.4. The following features were extracted from the z-axis (dorsoventral) and x-axis (head-to-foot) SCG ensemble averages from every location for all subjects: first and second maxima locations (0-250 ms), first and second maxima width (0-250 ms), first and second minima location (0-250 ms), first and second minima width (0-250 ms), first maximum location (250-500 ms), first maximum width (250-500 ms), first minimum location (250-500 ms), first minimum width (250-500 ms). Thus, a total of 12 features per axis for each sensor placement were extracted. We picked timing features from the early portion of the signal (0-250 ms) in order to captures events related to the aortic valve opening. Rather

than calculating only the R-AO interval, we explored a number of peaks timings and widths as these timings might also be related to cardiac mechanics. Additionally, although it is the timing features from the systolic portion of the signals that are related to PEP, we decided to explore the diastolic portion as well by adding peak timings and widths from the later portion of the signals (250-500 ms). The features are illustrated in Figure 20(b).

4.2.4 Regression Model

To estimate the PEP using the features extracted from the SCG signals, we trained a regression model. For every accelerometer placement, M features extracted from N ensemble averages were placed in an $N \times M$ matrix $\mathbf{X}$ while the corresponding PEP values are placed in an $N \times 1$ vector $\mathbf{y}_{PEP}$. These were then used to train a regression model which learns the relationship between $\mathbf{X}$ and $\mathbf{y}_{PEP}$. The learned model can then be used to estimate PEP for new heartbeats, given the features extracted from the SCG signals.

Many previous studies that utilize SCG signals to estimate hemodynamic parameters use linear regression to relate the SCG features to the estimated parameter [35, 41, 60, 66]. However, it is possible that the relationship between PEP and SCG features is not necessary linear, as is often the case in real data sets [67]. Therefore, rather than using only linear techniques, we chose to exploit recent advances in the field of machine learning: specifically, in this work, we used Extreme Gradient Boosting (XGBoost) regression which is a relatively new algorithm that has recently gained popularity as it is a computationally efficient implementation of a powerful ensemble learning technique [68]. XGBoost is an implementation of the gradient boosting machine learning algorithm [69]. This falls under a category of learning algorithms called ensemble methods, which combines multitudes of

estimators to predict a variable rather than using a single estimator [70]. XGBoost trains many regression trees iteratively, where new models predict the residual errors of previous ones and are then added together.

Specifically, let $\mathbf{x}$ be a vector containing all features extracted from one ensemble in our dataset. Let $h(\mathbf{x})$ denote a regression tree, which partitions the feature space into pieces and assigns a constant to each partition [71]. XGBoost trains many regression trees $h_m(\mathbf{x})$ in a consecutive manner, and adds them up such that the cumulative model $f_m(\mathbf{x})$ at iteration *m* is:

$$f_m(\mathbf{x}) = f_{m-1}(\mathbf{x}) + v h_m(\mathbf{x}) \tag{11}$$

$h_m(\mathbf{x})$ is trained to predict the error residuals between $f_m(\mathbf{x})$ and the target variable. The hyperparameter *v* is referred to as the learning rate and has the effect of shrinking the contribution of each individual tree $h_m(\mathbf{x})$, thus reducing the risk of overfitting [69]. Each regression tree is trained on a randomly sampled subset of the ensembles (rows of $\mathbf{X}$) and features (columns of $\mathbf{X}$). This introduces two hyper-parameters which are the factors by which the rows and columns of $\mathbf{X}$ are sampled.

Furthermore, while training the regression tree $h_m(\mathbf{x})$, the weights assigned to each partition of the feature space are L_2 regularized using a regularization hyper-parameter λ. Row and column subsampling as well as regularization has been reported to increase model accuracy and decrease overfitting [68-71]. The final regression model $f_{M_b}(\mathbf{x})$ emerges after M_b stages of iteratively applying equation (11), where M_b is referred to as the number of boosting rounds.

We used XGBoost regression to estimate PEP using SCG signals from different sensor locations. Estimated PEP results from different sensor locations were compared using a variation of the repeated cross-validation model assessment method discussed in [72].

Given that our dataset consists of 10 subjects, we first randomly paired these subjects into five groups. We then performed cross-validation by leaving one group (two subjects) out at each fold. At each fold of the cross-validation, we trained an XGBoost regressor on the data from all subjects except the ones that were left out. We predicted PEP for the left-out subjects and repeated this four more times to have PEP predictions for all ensembles from all subjects. We then calculated the root mean squared error (RMSE) between the estimated and ground truth PEPs. The entire process was repeated 50 times with a new random pairing of subjects each time. We calculated the cross-validation RMSE as the average of the RMSE scores from 50 repetitions. This approach was repeated for different SCG sensor locations and compared the resulting RMSE scores. We also combined pairs of sensor locations in order to explore which combination of two locations yields better PEP estimates. This was performed by combining features from a pair / multiple combinations of sensor locations and running the cross-validation procedure for different pairs / combinations of locations. We repeated this procedure using features only from the z-axis (dorsoventral), and both z+x axes combined (dorsoventral + head-to-foot). Single and multiple sensor comparisons were supported using statistical analysis of the cross-validation results which are described below in section 4.2.10. Using this approach, a global model was trained rather than multiple subject-specific models.

We performed this cross-validation procedure rather than using a leave one subject out cross-validation (LOSO-CV) to minimize the risk of overfitting the model to the dataset [73]. In LOSO-CV, only one subject is left out at each fold. Therefore, the training sets at each fold do not differ significantly. Leaving more subjects out leads to more variety in the training sets at each fold, and thus improves the generalizability of the resultant regression model. Additionally, there are many ways in which subjects can be grouped into pairs for leave-two-subjects-out cross-validation. To mitigate dependency on this grouping, we repeat the cross-validation 50 times with different pairing at each repetition.

4.2.5 SCG Sensor Location Comparison

To compare different SCG sensor locations, we trained XGBoost regressors on SCG features acquired from the different locations on the body. We assessed the ability of a sensor location in estimating PEP by calculating the (RMSE) between the estimated PEP values $(\hat{PEP})_i$ and the ground truth PEP acquired from the ICG signals (PEP_i):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left((\hat{PEP})_i - PEP_i\right)^2} \tag{12}$$

where N is the number of ensembles. Figure 20(c) shows a high-level block diagram of the process.

All XGBoost regressors trained for sensor location and axis comparison have the following hyper-parameter settings: learning rate=0.1, number of boosting rounds=200, column sampling factor=0.5, row sampling factor=0.5, regularization parameter (λ)=1.

These parameters were selected heuristically and the sensitivity of our results to them is analyzed in section 4.3.4.

4.2.6 Rigid vs. Flexible Interfacing Material Between the SCG Sensor and Sternum

To compare different methods of interfacing the SCG sensor to the sternum, we calculated the cross-validation RMSE scores by repeating the procedure explained in section 4.2.4 on signals acquired by placing the accelerometer (1) directly on the sternum; (2) placing two accelerometers on the upper and lower sternum using a rigid custom sensor housing; and (3) placing flexible material between the accelerometer and the skin. The process was repeated using only the z axis of the sensors and combining both z+x axes. For the rigid custom housing, we compared PEP estimates obtained using only the accelerometer placed on the upper sternum, only the lower sternum and combining both locations. Comparisons of different sensor interfacing methods were supported using statistical analysis as explain in section 4.2.10. The SCG signals obtained from the different interfaces are shown in Figure 21(b).

4.2.7 Feature Importance Evaluation

While evaluating different SCG sensor locations and interfacing methods, we trained XGBoost regressors using many features acquired from one or multiple sensors and axes as shown in Figure 20(b). However, only a portion of these features is relevant. The XGBoost regressor (like any other gradient boosting tree) trains regression trees which can be used to rank the features according to importance. Typically, the deep nodes of a tree divide using less important features while the main (first) node divides on the most important feature. The importance of features obtained from all trees are averaged resulting

in the final relative feature importance scores, which can then be used to rank the features. Mathematical details of the feature importance scoring and ranking are explained in detail in [71].

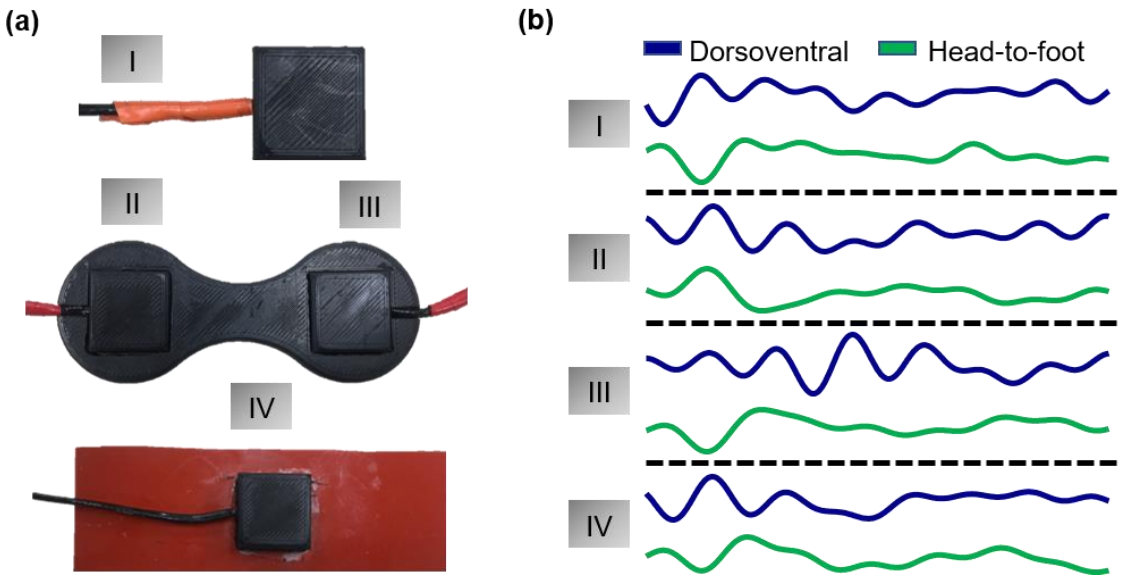

Figure 21. (a) Different methods of interfacing the ADXL354 accelerometer with the sternum. (b) SCG signals obtained from each of the different interfacing materials.

In order to evaluate which features generated from the SCG signals were more relevant in estimating PEP, we trained an XGBoost regressor on the pair of sensor combination and axis that gave the best PEP estimates (left clavicle + sternum, z-axis). All data from every subject were used to train the regressor and the resulting model was used to generate relative feature importance scores as described above. It should be noted that no testing set is required to score feature importance as we are not evaluating generalizability of the model for this portion of the study.

4.2.8 Comparing Different Regression Techniques

We hypothesized that XGBoost regression would perform better than linear regression models as well as other non-linear regression models. To address our hypothesis, we compared results obtained using SCG sensors (z-axis only) placed on the left clavicle and sternum (sensor placement combination with lowest cross-validated RMSE) and XGBoost regression to the same sensor placement combination (same feature set) but using simple multiple linear regression. This method uses ordinary least squares [71] and is the regressor of choice for previous related work [35, 41, 60, 66].

An extension of ordinary least squares regression is Ridge regression which penalizes regression coefficients by utilizing L_2 regularization to reduce overfitting [74]. This regression model is still linear but the L_2 regularization results in some coefficients to shrink, where the amount of shrinkage is controlled by a parameter α_{ridge}. Setting $\alpha_{ridge}=0$ is identical to ordinary least squares regression and coefficients are more heavily shrunk as α_{ridge} is increased. We assessed the performance of this regression technique as well on our dataset by training Ridge regression models with α_{ridge} ranging from 10^{-3} to 10^{2} logarithmically, keeping the feature and data set the same as explained above.

Another variant of ordinary least squares regression is Lasso regression which uses L_1 regularization to shrink regression coefficients via a regularization parameter α_{Lasso} [75]. Compared to Ridge regression (which utilizes L_2 regularization), Lasso tends to produce sparser linear models where many coefficients can be reduced to zero, which results in a form of feature selection. We also trained Lasso regressors on the same dataset varying α_{lasso} logarithmically from 10^{-3} to 10^{2}.

XGBoost is an ensemble method combining many base estimators [68]. Ensemble methods using regression tree estimators can fit complicated non-linear functions robustly, which might result in better estimations compared to linear models such as ordinary, Ridge, or Lasso regression. We compared XGBoost to two other ensemble regression methods: random forest regression and Extra-Trees regression.

Random forest regression like XGBoost trains many regression trees but unlike gradient boosting, each trained tree is independent. Trees are trained on a sample drawn with replacement and on a random subset of features [76]. Extra-Trees regression is similar to Random forest regression but adds more randomness to the model.

While Random forest tree nodes are divided using the most discriminative threshold, in Extra-Trees, a subset of thresholds is chosen for each feature and the best combination of these random splits are chosen [77]. We assessed the performance of Random forest and Extra-Trees regression on our dataset where regressors contain 200 trees and column sampling factor is chosen as 0.5 similar to XGBoost model parameters. Each tree was trained on a subset of features consisting of $\sqrt{n_{features}}$ features.

We compared the cross-validated RMSE resulting from the different regression models using the cross-validation procedure explained in section 4.2.4. The features and dataset were kept the same and only the regression technique was altered.

4.2.9 Evaluating the Effect of XGBoost Hyperparameters

Our choice of XGBoost hyperparameters as explained in section 4.2.4 was heuristic; however, variations of these parameters may alter results. An algorithm can be

considered robust if changes in hyper-parameters do not alter the results greatly. Contrarily, if small perturbations in parameters cause large variations in results, the algorithm might not generalize well due to sensitive dependence on the choice of parameters.

We evaluated the dependence of results on the learning rate by varying this parameter on a logarithmic grid of 50 points ranging from 10^{-2} to 10^{0}. We evaluated the cross-validated RMSE as explained in section 4.2.4 on the feature set derived from the SCG signals acquired from the left clavicle and sternum (z-axis), for each of 50 different learning rates, keeping everything else constant. We repeated this for 50 column sampling factor values on a linear grid from 0.1 to 1.0, 50 row sampling factor values on the same grid, and 50 regularization parameters (λ) on a logarithmic grid ranging from 10^{-2} to 10^{1}. While varying each parameter, we set all other parameters constant at the values described in Section 4.2.4.

4.2.10 Statistical Analysis

We performed statistical analysis on the cross-validated RMSE results to compare various SCG sensor locations. We performed leave two subjects out cross-validation 50 times and calculated RMSE for each repetition. In each of the 50 repetitions, the subjects were paired randomly, and this process was repeated for different sensor locations and combinations of locations. The random seed was fixed so that in the j^{th} repetition, the subjects were paired in the same way for all sensor locations/combinations. RMSE results from the 50 repetitions were compared using multiple comparison statistical testing for different locations and combinations. The Friedman Test was used to detect if any differences exist. For post-hoc testing, Wilcoxon signed rank test was performed on pairs

of sensor locations / combinations to be compared. Benjamini-Hochberg correction for multiple comparison was performed on the p-value from the post-hoc testing. These statistical tests and the reasons they should be used to compare machine learning models are discussed in detail in [78]. A similar procedure was followed to compare different sensor interfacing methods statistically. P-values below 0.05 were considered statistically significant.

4.3 Results

4.3.1 Comparison of Different Sensor Locations

Results comparing RMSE in milli-seconds (ms) from the SCG signals obtained from accelerometers placed on sternum, PMI, below the left clavicle and below the right clavicle, and every combination of these accelerometers for both the dorsoventral (z-axis) and the head-to-foot combined with the dorsoventral axis (z+x) are shown in Table 7. Statistically significant differences exist in these results according to Friedman test ($P<0.05$). To investigate where the significance exists, Wilcoxon signed rank test was performed on the different pairs of locations/combination of locations.

When comparing single locations, results showed that the z-axis of the signals from the sensors placed below the left and right clavicle provided the least RMSE in PEP estimates with 13.4±0.4 ms and 13.2±0.4 ms respectively (Figure 22(a), $P<0.05$ according to the Wlicoxon signed rank test comparing RC and LC to sternum and PMI). Additionally, features from head-to-foot SCG did not add substantial information to dorsoventral SCG features for all the signals except LC SCG, whose RMSE improves from 13.4±0.4 ms to

12.4±0.5 ms by combining the dorsoventral and head-to-foot features (P<0.05 according to Wilcoxon signed rank).

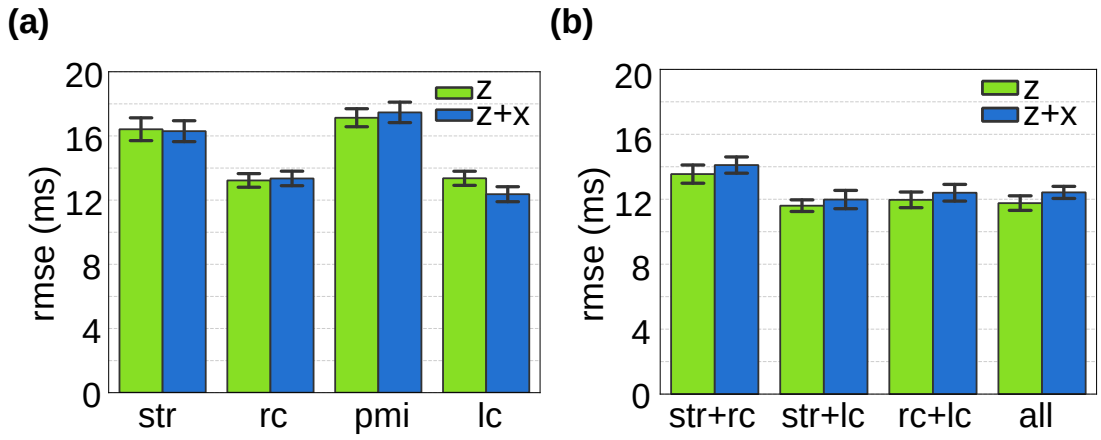

Figure 22. (a). RMSE from PEP estimated from features obtained from accelerometers placed on the sternum (Str), point of maximal impulse (PMI), below the left clavicle (LC), and below the right clavicle (RC) for both the dorsoventral axis (z-axis) and head-to-foot and dorsoventral axes combined (z+x axes). (b) RMSE from PEP estimated from features obtained from the best performing combination of accelerometer locations. It can be observed that adding more sensors does not substantially reduce the error obtained using one sensor below the left or right clavicle.

When combining features from signals from multiple locations, the best performing combination was the dorsoventral SCG from the sternum and LC with RMSE=11.6±0.4 ms (P<0.05 with every single location and combination except the combination of all sensors z-axis, Wilcoxon signed rank). Adding features from other locations and axes to the z-axis of these two locations did not improve the estimation (Figure 22(b)).

This best combination of signals (sternum+LC z-axis) was used in the feature importance analysis as described in Section 4.2.7. The results show that the timing of the first maximum of the dorsorventral LC SCG signal (0-250 ms) is the most important feature used in the XGBoost trees. Out of the top 15 features, nine features belonged to the LC

SCG signals and six belonged to the sternum SCG signal. The top 15 features are shown in Figure 23.

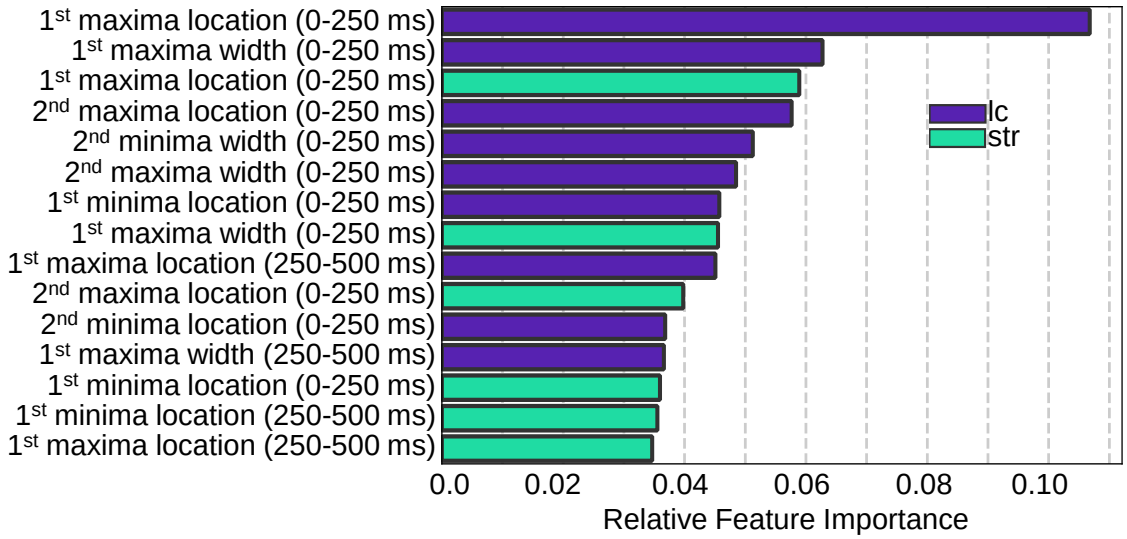

Figure 23. Ranking of best 15 features obtained from the combination of sensors and axis that rendered the lowest RMSE (Str+LC axis z).

Table 7. RMSE (ms) for PEP estimates from SCG signals measured from sensors placed in different locations (Str, PMI, LC, RC)

Location	z-axis RMSE	z+x axes RMSE
Str	16.42±0.71	16.30±0.65
PMI	17.13±0.56	17.47±0.64
LC	13.36±0.44	12.37±0.47
RC	13.23±0.42	13.35±0.45
Str+PMI	14.61±0.51	15.29±0.61
Str+LC	11.60±0.36	11.98±0.56
Str+RC	13.55±0.56	14.10±0.50
PMI+LC	13.12±0.49	12.57±0.46
PMI+RC	13.01±0.31	13.85±0.46
LC+RC	11.96±48	12.40±0.51
Str+PMI+RC	13.27±0.40	13.69±0.48
Str+PMI+LC	12.04±0.38	11.93±0.28
Str+RC+LC	11.97±0.42	12.37±0.46
PMI+RC+LC	12.19±0.50	12.84±0.49
Str+PMI+LC+RC	11.76±0.45	12.42±0.37

4.3.2 Comparison of Different Sensor Interfacing Material

Results comparing RMSE in ms from SCG signals obtained from accelerometers placed on the sternum directly (Str), using a flexible silicone rubber sheet (FStr), and using a rigid plastic mold coupling two accelerometers on upper and lower sternum (US and LS) are shown in Figure 24. Statistically significant differences exist in these results according to Friedman test ($P<0.05$). To investigate where the significance exists, Wilcoxon signed rank test was performed on the different pairs of interfacing methods.

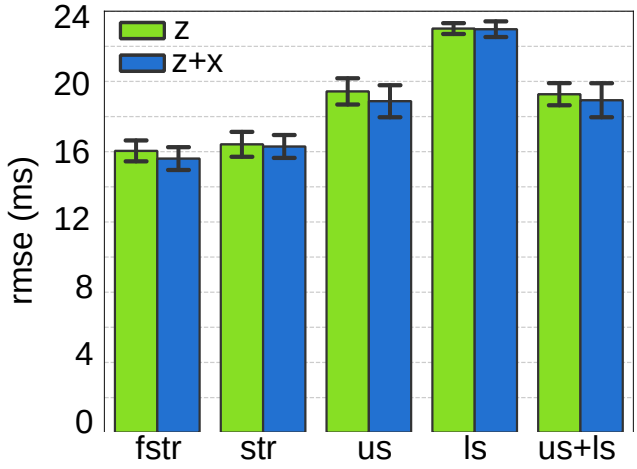

Figure 24. RMSE from PEP estimated from accelerometers placed on the sternum with different interfacing techniques: directly on the sternum (Str), in the middle of a silicone rubber sheet placed along on the sternum (fstr) and two accelerometers coupled by a rigid plastic mold and placed on the upper sternum (US) and lower sternum (LS)

Placing the accelerometer on the flexible silicone rubber sheet rather than directly on the sternum did not reduce the accuracy of the PEP estimates obtained from the SCG signals compared to placing the accelerometer directly on the sternum and even slightly improved it (16.0±0.6 ms vs. 16.4±0.7 ms respectively for the z-axis, $P<0.05$, Wilcoxon signed rank test; and 15.6±0.6 ms vs. 16.3±0.6 ms respectively for the z+x axes, $P<0.05$, Wilcoxon signed rank test). On the other hand, coupling the accelerometers using a rigid mold significantly reduced the accuracy of the PEP estimates obtained from these signals

even when features from both accelerometers placed on the upper and lower sternum were combined (P<0.05 when compared to str and fstr, Wilcoxon singed rank test).

4.3.3 Comparison of Different Regressors

When different regression techniques were compared keeping the feature set the same, it was observed that XGBoost produced the lowest RMSE as expected (Figure 25, P<0.05 according to Friedman test and P<0.05 when comparing XGBoost to all other regressors, Wilcoxon singed rank test). Compared to XGBoost, ordinary least squares regression resulted in an RMSE that was 8 ms higher. Introduction of L_2 regularization via Ridge regression does not improve linear regression results. Figure 25 shows results only for α_{ridge}=1, but results did not alter substantially for the range of α_{ridge} values tested. L_1 regularization via Lasso regression improves linear regression RMSE results by around 2 ms (P<0.05, Wilcoxon singed rank test), however, this is still substantially greater error than corresponding RMSE results using XGBoost.

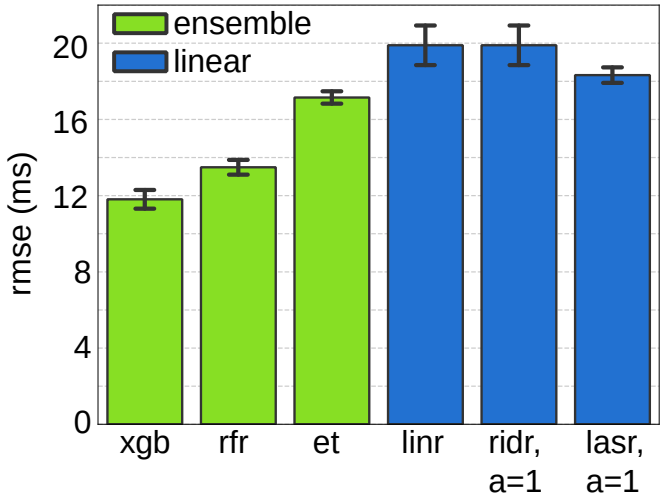

Figure 25. Comparing RMSE for PEP estimates obtained using ensemble regression models vs. linear regression models on features obtained from SCG signals that performed best with XGBoost (i.e., LC+sternum z-axis).

In Figure 25, ensemble learning methods have been highlighted in green while linear methods are shown in blue. As hypothesized, ensemble methods produce lower RMSE when compared to generalized linear methods. Out of the ensemble methods, we find that random forest regression performs better than Extra-Trees regression ($P<0.05$, Wilcoxon signed rank test), while XGBoost maintains the best performance.

4.3.4 Effect of XGBoost Hyperparameters

Figure 26(a) shows the variation in cross-validated RMSE as the learning rate is varied. The variations in RMSE for learning rates in the range 2×10^{-2} to 4×10^{-1} is minimal. However, learning rate must not be increased or decreased too much, as values above 0.4 and below 2×10^{-2} lead to deterioration in RMSE.

Figure 26(b) shows the effect of varying the column sampling factor on the RMSE. It can be seen that for values between 0.4 and 1.0, the variation in RMSE is minimal but for values below 0.4, RMSE deteriorates.

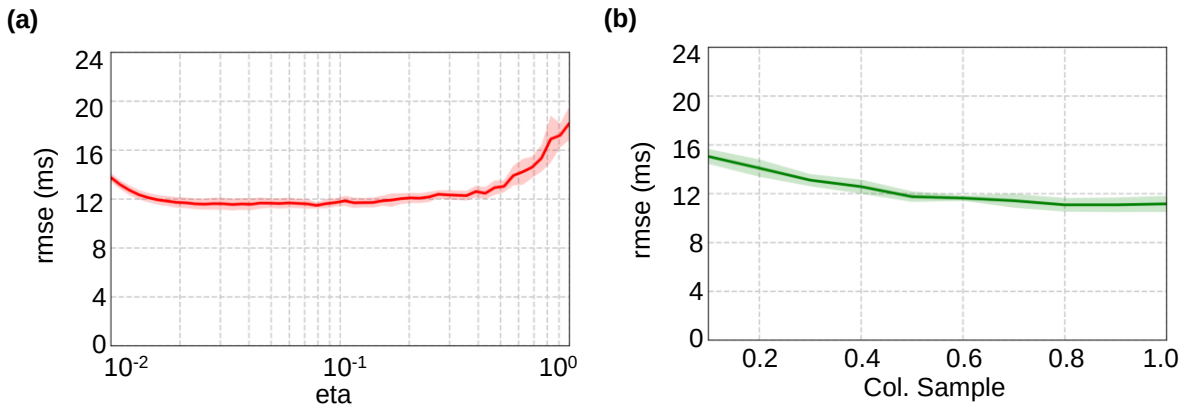

Figure 26. (a) RMSE for PEP estimates obtained using XGBoost on features obtained from LC+sternum z-axis while varying the learning rate parameter (b) RMSE for PEP estimates obtained using XGBoost on features obtained from LC+sternum z-axis while varying the column sample parameter.

We also investigated the effect of varying the row sampling factor and the regularization parameter on RMSE as discussed in section 4.2.9. Varying these parameters within the ranges explained in section 4.2.9 lead to nearly no changes in RMSE (<1 ms change in RMSE).

4.4 Discussion

Our universal regression model showed, for the first time, that SCG signals collected from below the left or right clavicle provide better PEP estimates than those collected from the sternum. To understand why LC and RC features perform better than sternum and PMI, we calculated Spearman correlation between the top 3 features from every location and the I-wave, J-peak, and K-point of the weighing scale BCG. The reason we perform this correlation is that it has been demonstrated that BCG, which represents total body displacement from the heart beats is a good alternative to ICG for measuring PEP [30]. Results showed no correlation between the sternum and PMI features with BCG features and good correlation between LC and RC features with BCG features, specifically J-peak and I-wave (Figure 27). Hence, it can be inferred that SCG signals obtained from below the left and right clavicle are more representative of total body displacement from the heart beats than SCG obtained from sternum and PMI.

Additionally, when examining the top features from the best combination of signals (sternum + LC), most top features were from LC SCG rather than from sternum SCG, which can be attributed to the fact that LC performs better than the sternum in terms of PEP estimation error for our dataset.

Beyond providing insight into sensor placement, the regression models differentiated the effects of different interfacing methods. Finding that SCG obtained from placing the sensor on a silicone rubber sheet performs as well, and even slightly better, in PEP estimation than SCG obtained from placing the sensor in direct contact with the skin could be an indication that the sensor can be worn over a thin layer of clothing (tightly fitted over the body to ensure good coupling of the sensor to the skin) without affecting the measurements. This is an important finding that supports the practicality of a wearable SCG device, and warrants further investigation.

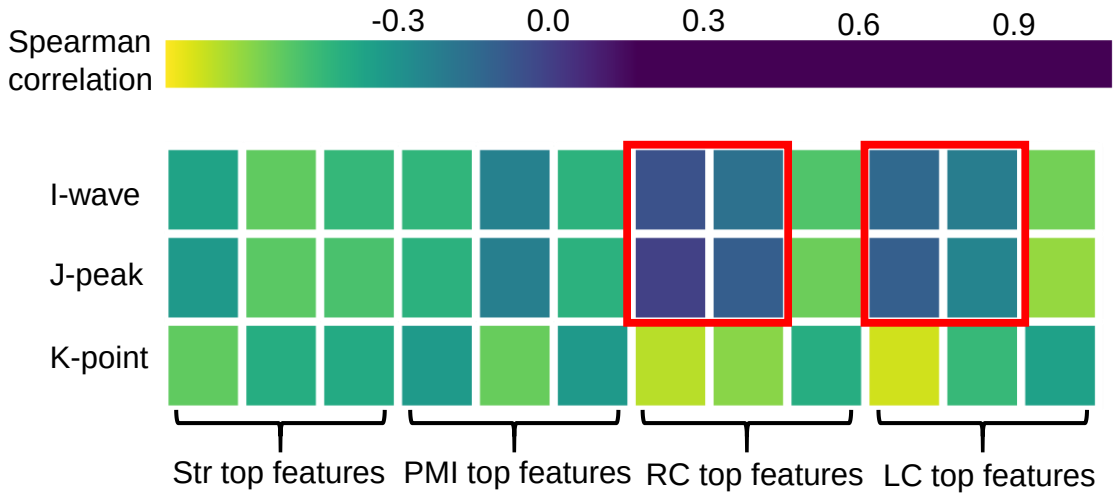

Figure 27. Spearman correlation between the top 3 features from signals from every location and the I-wave, J-peak, and K-point of scale BCG. There was almost no correlation between sternum and PMI features with BCG features while good correlation was obtained between RC and LC features with BCG features and specifically the I-wave and J-peak. (highlighted with the red boxes).

Furthermore, we demonstrated in this paper that the relation between the extracted features and PEP is better characterized by non-linear models rather than linear ones. Ensemble methods, were shown to model non-linear relationships between the predictors and PEP well [70]. Ensemble methods also attempt to perform automatic feature selection by picking random subsets of the whole feature set which might be more effective than

intrinsic feature selection using L_1 or L_2 regularization [71, 76]. In general, ensemble learners produce powerful estimators as they combine a diverse set of models. Each regression tree in the ensemble is built by randomly subsampling instances and features. A committee of these models is robust (reduced variance) and generalizes well [75].

From the examined linear regression methods, introducing L_1 regularization via Lasso regression slightly improved the linear regression while L_2 regularization via Ridge regression did not. This can be attributed to the fact that there is a large number of features and L_1 regularization produces better intrinsic feature selection (via shrinkage) than an L_2 penalty. On the other hand, from the examined ensemble regression methods, Extra-Trees regression performing worse than random forest regression can be linked to the increased randomness in determining tree node division thresholds in Extra-Trees which might account for the degradation in performance. Finally, XGBoost performed best, possibly because each tree in the ensemble is iteratively trained according to the prediction errors of previous trees, whereas the other methods train independent trees. XGBoost also regularizes each tree in the ensemble to improve generalization and reduce overfitting, which is not performed in random forests or Extra-Trees.

A drawback of XGBoost compared to the other techniques considered is the abundance of hyperparameters to tune. However, the performance was found to be robust to changes in hyperparameters as variations in the regularization parameter row sampling factor, and in learning rate even by an order of magnitude caused infinitesimal variation in RMSE. Large values of the learning rate mean that each tree in the gradient boosting ensemble contribute more to the model which reduces the regularization effect, rendering the trained model more prone to overfitting. On the other hand, as the learning rate is

decreased, each tree's contribution decreases and in the extreme of a null learning rate, the consecutive trees do not contribute to the mode at all, meaning the gradient boosting is no longer performed. Therefore, a learning rate value > 0.4 or < 2×10^{-2} resulted in deterioration in RMSE. Finally, when varying the column sampling factor, values below 0.4 worsened the RMSE, because as the column sampling factor decreases, increasingly smaller subsets of the features space are used to train each tree in the ensemble.

In general, our method was robust to changes in hyperparameters, and thus we did not perform hyperparameter tuning. While hyperparameter tuning via extensive grid search or random search can likely lead to improvement in RMSE, this approach would also be very time consuming and computationally expensive. Robustness to variations in the hyperparameters also means that the gains due to tuning might be relatively small compared to gains due to better sensor positioning.

4.5 Conclusions and Future Work

In this chapter, we developed universal regression models to estimate PEP from SCG signals measured from multiple locations and compared the outputs of the resulting regression models by calculating the RMSE between the estimated PEP values with the ground truths PEP values obtained from ICG. We demonstrated that ensemble regression models provide significantly more accurate PEP estimates than linear regression models. Additionally, in our dataset, we showed, for the first time, that placing an accelerometer below the right or left clavicle results in better PEP estimates than placing the accelerometer on the sternum, which is the common placement in SCG literature. We also found that SCG signals obtained from below the left and right clavicle are more

representative of total body displacement from heart beats as their features were better correlated with BCG features than sternum SCG. Finally, we showed that placing the sensor on a flexible silicon rubber sheet rather than directly on the skin does not reduce the accuracy of the measurements which could be an indication that the wearable SCG sensing device can be placed over one's clothes without deteriorating its performance.

Future work should include analyzing SCG signals obtained from placing the sensor on different fabrics to verify that the SCG wearable device can be placed over clothes. More locations should also be investigated to understand the relation of whole body vibrations to SCG. More subjects should be included in future studies to be able to leave more subjects out during the model training and hyperparameter tuning should be performed to minimize RMSE. Finally, the same approach used in this paper can be used to estimate other cardiac parameters such as left ventricular ejection time (LVET) which with PEP can provide us with a measure of changes in cardiac contractility.

CHAPTER 5. CONCLUSION AND FUTURE DIRECTION

5.1 Conclusion

The American Heart Association forecasts that by 2030, there will be a 46% increase in the number of HF patients and a 127% in healthcare costs associated with the diseases from today's estimates. Consequently, it can be deduced that this increase will result in a shortage in the number of healthcare providers per patient. Therefore, there is a compelling need to transfer care from clinics to the homes of patients to increase accessibility and decrease the overall cost of care. This would allow care to be automatically optimized to the needs of the patient with proactive, feedback-controlled therapies, and would thereby reduce the burden on the healthcare system by reducing the need for costly reactive measures such as emergency room visits.

The methods and algorithms developed and discussed in this work improve the robustness of STIs measured using BCG and SCG signals and therefore significantly improve out-of-clinic monitoring of heart failure patients. This work can also advance the understanding of basic physiology by providing a platform for easy, unobtrusive, and in some cases continuous, measurements of the mechanical parameters of the cardiovascular system. A lot of active research is ongoing with BCG and SCG as unobtrusive alternatives to echocardiography and ICG for monitoring the mechanical function of the heart. However, it is important to ensure that the research can cross from the experimental clinical to non-clinical domains while maintaining the same level of reliability. This was the motivation behind this work as this dissertation addresses many of the limitations preventing this transition into the non-clinical, uncontrolled settings through the use of both

novel algorithms and measurement modalities that can potentially improve the accuracy of cardiovascular parameters measured outside the clinics.

5.2 Future Directions

Various potential future research endeavors can stem from this research. First, working on developing affordable high bandwidth systems, with similar performance to the force plate used in this research, would enable the use of these systems at homes rather than in public places such as health centers and grocery stores. This would in turn facilitate the monitoring process by guaranteeing the patient's ability to obtain multiple measurements throughout the day.

The models and algorithms developed in chapters 3 and 4 to detect misplacement and universally estimate PEP can be extended to other hemodynamic parameters, such as the left ventricular ejection time (LVET), which, with PEP, can provide a measure of changes in cardiac contractility. Additionally, these algorithms take into consideration the different physiological states the person may pass through during the day, but they do not consider the pathophysiological differences such as the changes that may occur in a patient's condition. Therefore, since the end goal is to use these algorithms on BCG / SCG signals to monitor HF patients, they should be tested on a population of those patients, both in hospitals and at homes, to make sure they provide an analogous level of performance.

Finally, a known problem with BCG and SCG signals is the large inter-subject variability of the waveforms, which is why these signals have not been used yet for diagnosis, but can used for monitoring the cardiac health of each person individually, due to the low intra-subject variability in the signals over serial measurements. However, after

developing the universal regression model discussed in chapter 4, it is possible that this would bypass the inter-subject variability problem and could be leveraged as a step to enabling early detection of HF using BCG / SCG signals. To test for that, controlled experiments need to be run on healthy subjects and HF patients to determine if a model can be designed to distinguish between the two categories and assess patients with HF.

REFERENCES

[1] S. Mangini, P. V. Pires, F. G. M. Braga, and F. Bacal, "Decompensated heart failure," *Einstein,* vol. 11, no. 3, pp. 383-391, Jul-Sep 02/15/received, 08/13/accepted 2013.

[2] S. A. Hunt *et al.*, "2009 Focused Update Incorporated Into the ACC/AHA 2005 Guidelines for the Diagnosis and Management of Heart Failure in AdultsA Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines Developed in Collaboration With the International Society for Heart and Lung Transplantation," *Journal of the American College of Cardiology,* vol. 53, no. 15, pp. e1-e90, 2009.

[3] A. H. Association, "What is Heart Failure?," ed, 2015.

[4] D. Mozaffarian *et al.*, "Executive Summary: Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association," *Circulation,* vol. 133, no. 4, p. 447, 2016.

[5] N. C. f. H. Statistics, "Mortality multiple cause micro-data files, 2011: public-use data file and documentation: NHLBI tabulations," ed, 2014.

[6] P. A. Heidenreich *et al.*, "Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association," *Circulation,* vol. 123, no. 8, pp. 933-44, Mar 1 2011.

[7] A. L. Bui and G. C. Fonarow, "Home monitoring for heart failure management," *J Am Coll Cardiol,* vol. 59, no. 2, pp. 97-104, Jan 10 2012.

[8] A. S. Desai and L. W. Stevenson, "Rehospitalization for heart failure predict or prevent?," *Circulation,* vol. 126, no. 4, pp. 501-506, 2012.

[9] A. Giordano *et al.*, "Multicenter randomised trial on home-based telemanagement to prevent hospital readmission of patients with chronic heart failure," *International journal of cardiology,* vol. 131, no. 2, pp. 192-199, 2009.

[10] S. I. Chaudhry, Y. Wang, J. Concato, T. M. Gill, and H. M. Krumholz, "Patterns of weight change preceding hospitalization for heart failure," *Circulation,* vol. 116, no. 14, pp. 1549-1554, 2007.

[11] S. I. Chaudhry *et al.*, "Telemonitoring in patients with heart failure," *New England Journal of Medicine,* vol. 363, no. 24, pp. 2301-2309, 2010.

[12] S. D. Anker *et al.*, "Prognostic importance of weight loss in chronic heart failure and the effect of treatment with angiotensin-converting-enzyme inhibitors: an observational study," *The Lancet,* vol. 361, no. 9363, pp. 1077-1083, 2003.

[13] E. E. Wolfel, "Can we predict and prevent the onset of acute decompensated heart failure?," *Circulation,* vol. 116, no. 14, pp. 1526-1529, 2007.

[14] P. B. Adamson *et al.*, "Ongoing right ventricular hemodynamics in heart failure: clinical value of measurements derived from an implantable monitoring system," *Journal of the American College of Cardiology,* vol. 41, no. 4, pp. 565-571, 2003.

[15] W. T. Abraham *et al.*, "Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial," *The Lancet,* vol. 377, no. 9766, pp. 658-666, 2011.

[16] R. P. Lewis, S. E. Rittogers, W. F. Froester, and H. Boudoulas, "A critical review of the systolic time intervals," *Circulation,* vol. 56, no. 2, pp. 146-58, Aug 1977.

[17] P. Abrahams, *How the Body Works*. New York: Metro Books, 2012.

[18] D. B. Newlin and R. W. Levenson, "Pre-ejection period: measuring beta-adrenergic influences upon the heart," *Psychophysiology,* vol. 16, no. 6, pp. 546-53, Nov 1979.

[19] R. C. Talley, J. F. Meyer, and J. L. McNay, "Evaluation of the pre-ejection period as an estimate of myocardial contractility in dogs," *The American Journal of Cardiology,* vol. 27, no. 4, pp. 384-391, 1971/04/01 1971.

[20] A. M. Katz, *Physiology of the heart*. New York: Raven Press, 1977, pp. xiii, 450 p.

[21] D. Dubin, *Rapid Interpretation of EKG's*. USA: Cover Publishing Company, 1996, 1996.

[22] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement,* vol. 28, no. 3, p. R1, 2007.

[23] A. Reisner, P. A. Shaltis, D. McCombie, and H. H. Asada, "Utility of the photoplethysmogram in circulatory monitoring," *The Journal of the American Society of Anesthesiologists,* vol. 108, no. 5, pp. 950-958, 2008.

[24] Y.-L. Zheng *et al.*, "Unobtrusive sensing and wearable devices for health informatics," *IEEE Transactions on Biomedical Engineering,* vol. 61, no. 5, pp. 1538-1554, 2014.

[25] M. Etemadi and O. T. Inan, "Wearable Ballistocardiogram and Seismocardiogram Systems for Health and Performance," *Journal of Applied Physiology,* p. jap. 00298.2017, 2017.

[26] M. Ulbrich *et al.*, "The IMPACT shirt: textile integrated and portable impedance cardiography," *Physiological measurement,* vol. 35, no. 6, p. 1181, 2014.

[27] R. Patterson, "Fundamentals of impedance cardiography," *IEEE Engineering in Medicine and Biology magazine,* vol. 8, no. 1, pp. 35-38, 1989.

[28] G. Cybulski, E. Michalak, E. Koźluk, A. Piątkowska, and W. Niewiadomski, "Stroke volume and systolic time intervals: Beat-to-beat comparison between echocardiography and ambulatory impedance cardiography in supine and tilted positions," *Medical and Biological Engineering and Computing*, vol. 42, no. 5, pp. 707-711, 2004// 2004.

[29] V. V. Ermishkin *et al.*, "Beat-by-beat changes in pre-ejection period during functional tests evaluated by impedance aortography: a step to a left ventricular contractility monitoring," in *13th International Conference on Electrical Bioimpedance and the 8th Conference on Electrical Impedance Tomography: ICEBI 2007, August 29th - September 2nd 2007, Graz, Austria*, H. Scharfetter and R. Merwa, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 655-658.

[30] M. Etemadi, O. T. Inan, L. Giovangrandi, and G. T. Kovacs, "Rapid assessment of cardiac contractility on a home bathroom scale," *IEEE Trans Inf Technol Biomed*, vol. 15, no. 6, pp. 864-9, Nov 2011.

[31] M. Packer *et al.*, "Utility of impedance cardiography for the identification of short-term risk of clinical decompensation in stable patients with chronic heart failure," *J Am Coll Cardiol*, vol. 47, no. 11, pp. 2245-52, Jun 6 2006.

[32] J. P. Piccini and P. Hranitzky, "Diagnostic monitoring strategies in heart failure management," *Am Heart J*, vol. 153, no. 4 Suppl, pp. 12-7, Apr 2007.

[33] O. T. Inan *et al.*, "Ballistocardiography and seismocardiography: a review of recent advances," *IEEE J Biomed Health Inform*, vol. 19, no. 4, pp. 1414-27, Jul 2015.

[34] I. Starr, A. J. Rawson, H. A. Schroeder, and N. R. Joseph, "Studies on the estimation of cardiac output in man, and of abnormalities in cardiac funtion, from the heart's recoil and the blood's impacts; the ballistocardiogram," *American Journal of Physiology*, vol. 127, no. 1939, pp. 1-28, 1939.

[35] K. Tavakolian, A. P. Blaber, B. Ngai, and B. Kaminska, "Estimation of hemodynamic parameters from seismocardiogram," in *2010 Computing in Cardiology*, 2010, pp. 1055-1058: IEEE.

[36] J. Gordon, "Certain molar movements of the human body produced by the circulation of the blood," *Journal of Anatomy and Physiology*, vol. 11, no. Pt 3, p. 533, 1877.

[37] A. Lindqvist, K. Pihlajamäki, J. Jalonen, V. Laaksonen, and J. Alihanka, "Static-charge-sensitive bed ballistocardiography in cardiovascular monitoring," *Clinical Physiology*, vol. 16, no. 1, pp. 23-30, 1996.

[38] H. J. Baek, G. S. Chung, K. K. Kim, and K. S. Park, "A smart health monitoring chair for nonintrusive measurement of biological signals," *IEEE transactions on Information Technology in Biomedicine*, vol. 16, no. 1, pp. 150-158, 2012.

[39] O. Inan, M. Etemadi, A. Paloma, L. Giovangrandi, and G. Kovacs, "Non-invasive cardiac output trending during exercise recovery on a bathroom-scale-based ballistocardiograph," *Physiological measurement,* vol. 30, no. 3, p. 261, 2009.

[40] K. Tavakolian, B. Ngai, A. P. Blaber, and B. Kaminska, "Infrasonic cardiac signals: Complementary windows to cardiovascular dynamics," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 4275-4278.

[41] P. Castiglioni, A. Faini, G. Parati, and M. Di Rienzo, "Wearable seismocardiography," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2007, pp. 3954-7, 2007.

[42] Y. Chuo *et al.,* "Mechanically flexible wireless multisensor platform for human physical activity and vitals monitoring," *IEEE Trans Biomed Circuits Syst,* vol. 4, no. 5, pp. 281-94, Oct 2010.

[43] M. Etemadi, O. T. Inan, J. A. Heller, S. Hersek, L. Klein, and S. Roy, "A Wearable Patch to Enable Long-Term Monitoring of Environmental, Activity and Hemodynamics Variables," *IEEE Trans Biomed Circuits Syst,* vol. 10, no. 2, pp. 280-8, Apr 2016.

[44] E. Pinheiro, O. Postolache, and P. Girão, "Theory and developments in an unobtrusive cardiovascular system representation: ballistocardiography," *The open biomedical engineering journal,* vol. 4, p. 201, 2010.

[45] M. D. Rienzo *et al.*, "A wearable system for the seismocardiogram assessment in daily life conditions," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 4263-4266.

[46] O. T. Inan *et al.*, "Using Ballistocardiography to Monitor Left Ventricular Function in Heart Failure Patients," *Journal of Cardiac Failure,* vol. 22, no. 8, p. S45.

[47] O. T. Inan, M. Etemadi, R. M. Wiard, L. Giovangrandi, and G. T. Kovacs, "Robust ballistocardiogram acquisition for home monitoring," *Physiol Meas,* vol. 30, no. 2, pp. 169-85, Feb 2009.

[48] A. Sherwood, M. T. Allen, J. Fahrenberg, R. M. Kelsey, W. R. Lovallo, and L. J. van Doornen, "Methodological guidelines for impedance cardiography," *Psychophysiology,* vol. 27, no. 1, pp. 1-23, Jan 1990.

[49] G. Cybulski and SpringerLink (Online service), Ambulatory impedance cardiography the systems and their applications, Berlin ; London: Springer,, 2010, p. p. [Online]. Available: SpringerLink http://dx.doi.org/10.1007/978-3-642-11987-3 MIT Access Only.

[50] J. Martin Bland and D. Altman, "STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL

MEASUREMENT," *The Lancet,* vol. 327, no. 8476, pp. 307-310, 1986/02/08 1986.

[51] T. Arai, K. Lee, and R. J. Cohen, "Cardiac output and stroke volume estimation using a hybrid of three Windkessel models," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010, pp. 4971-4974.

[52] R. Chatti *et al.*, "Comparison of two versions of the Vigileo-FloTrac™ system (1.03 and 1.07) for stroke volume estimation: a multicentre, blinded comparison with oesophageal Doppler measurements," *British Journal of Anaesthesia,* vol. 102, no. 4, pp. 463-469, April 1, 2009 2009.

[53] A. F. Rabbi, R. Gagarin, G. C. Huang, and M. F. Iskander, "Stroke volume estimation from the new noninvasive cardiopulmonary stethoscope," in *2014 IEEE Antennas and Propagation Society International Symposium (APSURSI)*, 2014, pp. 967-968.

[54] C. Schmidt *et al.*, "Comparison of electrical velocimetry and transoesophageal Doppler echocardiography for measuring stroke volume and cardiac output," *Br J Anaesth,* vol. 95, no. 5, pp. 603-10, Nov 2005.

[55] L. A. H. Critchley and J. A. J. H. Critchley, "A Meta-Analysis of Studies Using Bias and Precision Statistics to Compare Cardiac Output Measurement Techniques," *Journal of Clinical Monitoring and Computing,* journal article vol. 15, no. 2, pp. 85-91, 1999.

[56] M. Di Rienzo, P. Lombardi, D. Scurati, and E. Vaini, "A new technological platform for the multisite assessment of 3D seismocardiogram and pulse transit time in cardiac patients," in *Computing in Cardiology*, Vancouver, BC, 2016.

[57] L. S. Evangelista, H. Ghasemzadeh, J.-A. Lee, R. Fallahzadeh, M. Sarrafzadeh, and D. K. Moser, "Predicting adherence to use of remote health monitoring systems in a cohort of patients with chronic heart failure," *Technology and Health Care*, no. Preprint, pp. 1-9, 2016.

[58] Anatomography, "Sternum Front," in *BodyParts3D/Anatomography*, ed, 2012.

[59] F. Khosrow-Khavar, K. Tavakolian, A. Blaber, and C. Menon, "Automatic and Robust Delineation of the Fiducial Points of the Seismocardiogram Signal for Non-invasive Estimation of Cardiac Time Intervals," *IEEE Transactions on Biomedical Engineering,* 2016.

[60] A. Q. Javaid *et al.*, "Quantifying and Reducing Motion Artifacts in Wearable Seismocardiogram Measurements during Walking to Assess Left Ventricular Health," *IEEE Transactions on Biomedical Engineering,* 2016.

[61] T. M. Mitchell, *Machine Learning*. New York: McGraw-Hill, 1997, pp. xvii, 414 p.

[62] P. L. Enright, "The six-minute walk test," *Respiratory care,* vol. 48, no. 8, pp. 783-785, 2003.

[63] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences,* vol. 55, no. 1, pp. 119-139, 1997/08/01 1997.

[64] J. R. Quinlan, *C4.5 : programs for machine learning* (Morgan Kaufmann series in machine learning). San Mateo, Calif.: Morgan Kaufmann Publishers, 1993, pp. x, 302 p.

[65] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *International Journal of Machine Learning Technology,* vol. 2, no. 1, pp. 37-63, 2011.

[66] K. Tavakolian *et al.*, "Myocardial contractility: A seismocardiography approach," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 3801-3804: IEEE.

[67] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.

[68] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785-794: ACM.

[69] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics,* pp. 1189-1232, 2001.

[70] T. G. Dietterich, "Ensemble learning," *The handbook of brain theory and neural networks,* vol. 2, pp. 110-125, 2002.

[71] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Springer series in statistics New York, 2001.

[72] A. Dupuy and R. M. Simon, "Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting," *Journal of the National Cancer Institute,* vol. 99, no. 2, pp. 147-157, 2007.

[73] S. Hersek *et al.*, "Wearable Vector Electrical Bioimpedance System to Assess Knee Joint Health," *IEEE Transactions on Biomedical Engineering,* 2016.

[74] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics,* vol. 12, no. 1, pp. 55-67, 1970.

[75] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267-288, 1996.

[76] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R news*, vol. 2, no. 3, pp. 18-22, 2002.

[77] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning*, vol. 63, no. 1, pp. 3-42, 2006.

[78] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine learning research*, vol. 7, no. Jan, pp. 1-30, 2006.